

FIGURE 1 - General Overview of Distributed File Storage System

communication with other server nodes

communication with other server nodes

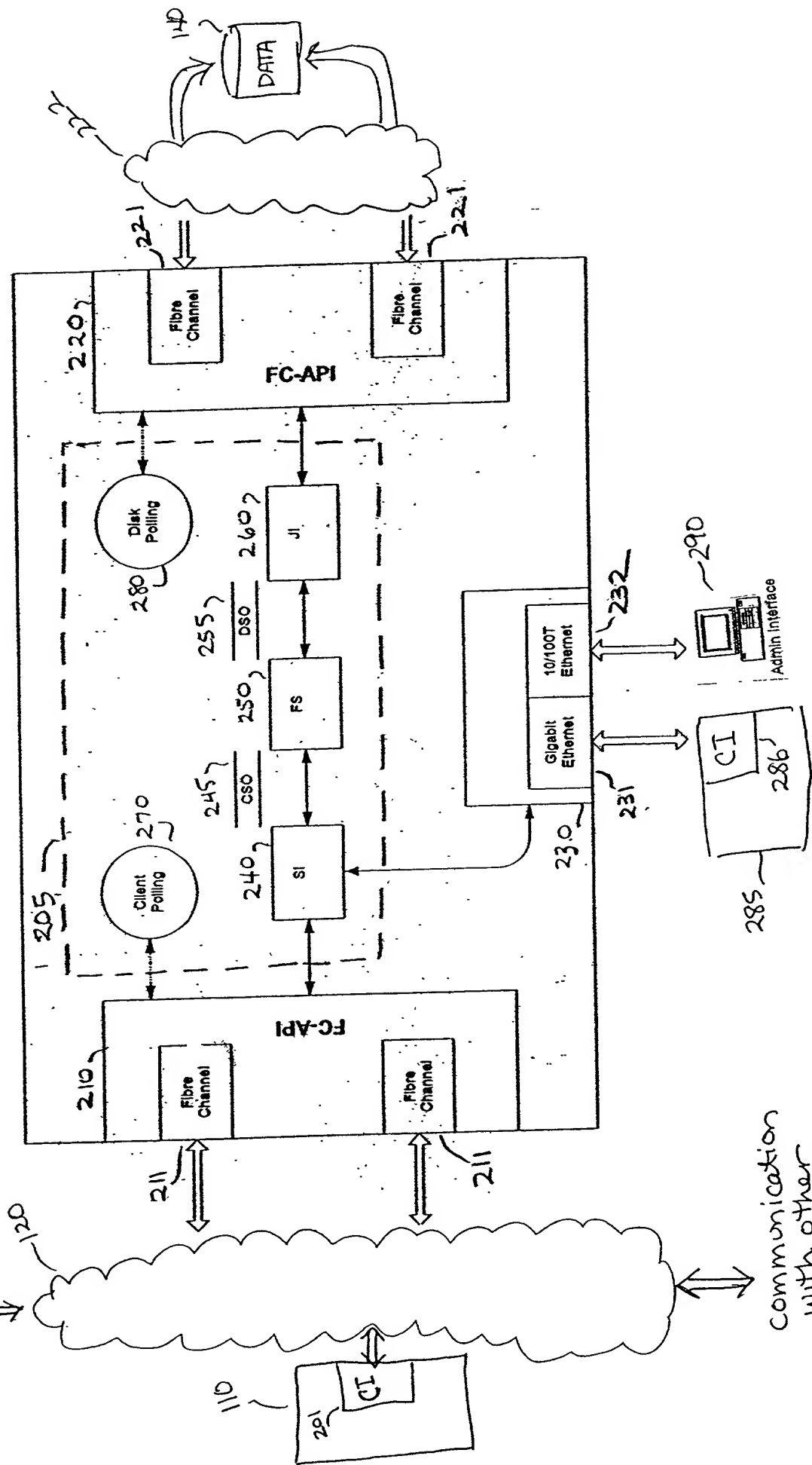


FIGURE 2 : One Embodiment of a Server Node

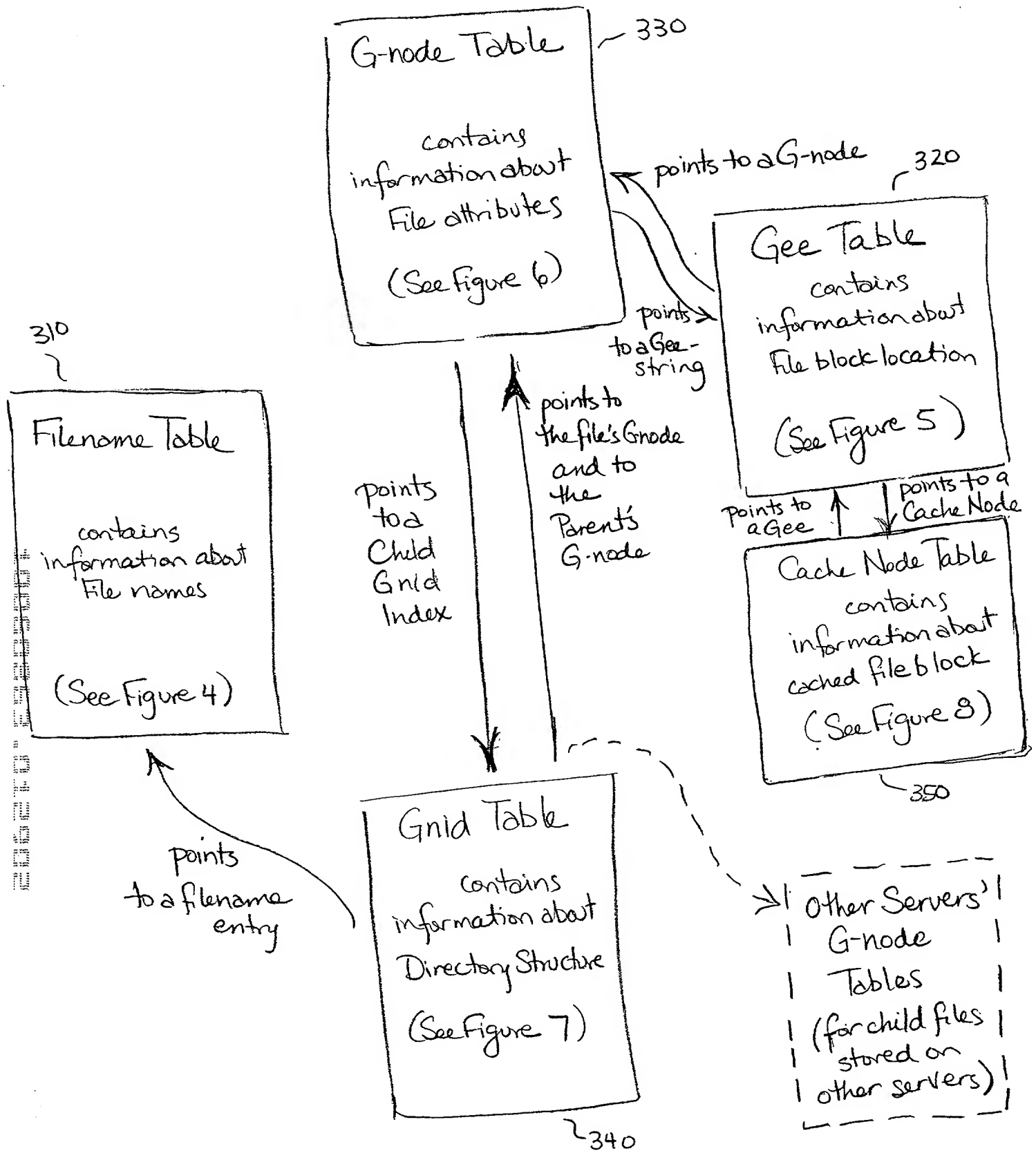


FIGURE 3 - Five metadata structures

310

		Array Index			Array Index	
410	411	70	<SOS>		85	<DS>
	412	71	csum		86	csum
	413	72	3		87	4
	414	73	'D'		88	'F'
		74	'o'		89	'r'
75		'e'		90	'o'	
420	421	76	<SOS>		91	'g'
	422	77	csum		92	<SOS>
	423	78	6		93	csum
	424	79	'T'		94	2
		80	'h'		95	'I'
		81	'o'		96	't'
		82	'm'		97	<SOS>
		83	'a'			
		84	'S'			

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

FIGURE 4 - Sample Portion of a Filename Table

320

590 591 592

	Index	G-Code	Data	File Logical Block
S10-	45	GNODE	Gnode = 67, Extent = 2, Root = TRUE	
S11-	46	DATA	Disk Logical Blocks: 456, 457 Drive 13	1
S12-	47	DATA	Disk Logical Blocks: 667, 668 Drive 15	2
S13-	48	DATA	Disk Logical Blocks: 112, 113 Drive 19	3
S14-	49	PARITY	Disk Logical Blocks: 554, 555 Drive 2	
S15-	50	DATA	Disk Logical Blocks: 458, 459 Drive 13	4
S16-	51	DATA	Disk Logical Blocks: 669, 670 Drive 15	5
S17-	52	DATA	Disk Logical Blocks: 119, 120 Drive 19	6
S18-	53	PARITY	Disk Logical Blocks: 556, 557 Drive 2	
S19-	54	LINK	Index 76	
	
S20-	76	GNODE	Gnode = 67, Extent = 3, Root = FALSE	
S21-	77	DATA	Disk Logical Blocks: 460, 461, 462 Drive 13	7
S22-	78	DATA	Disk Logical Blocks: 671, 672, 673 Drive 15	8
S23-	79	PARITY	Disk Logical Blocks: 121, 122, 123 Drive 19	
S24-	80	LINK	Index 88	
	
S25-	88	GNODE	Gnode = 67, Extent = 3, Root = FALSE	
S26-	89	DATA	Disk Logical Blocks: 463, 464, 465 Drive 13	9
S27-	90	DATA	Disk Logical Blocks: 674, 675, 676 Drive 15	10
S28-	91	PARITY	Disk Logical Blocks: 124, 125, 126 Drive 19	
S29-	92	GNODE	Gnode = 43, Extent = 4, Root = FALSE	
	

550
551
552

500

FIGURE 5 - Sample Portion of a Gee Table

00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000

Attribute Data	
602	File Attribute - type
604	File Attribute - mode
606	File Attribute - links
608	File Attribute - uid
610	File Attribute - gid
612	File Attribute - size
614	File Attribute - used
620	File Attribute - fileId
622	File Attribute - atime
624	File Attribute - mtime
626	File Attribute - ctime
628	Child Gnid Index
630	Gee Index - Last Used
631	Gee Offset - Last Used
632	Gee Index - Midpoint
633	Gee Offset - Midpoint
634	Gee Index - Tail
635	Gee Offset - Tail
636	Gee Index - Root
638	Gnode Status
640	Quick Shot Status
642	Quick Shot Link

600

FIGURE 6 - G-NODE ATTRIBUTES

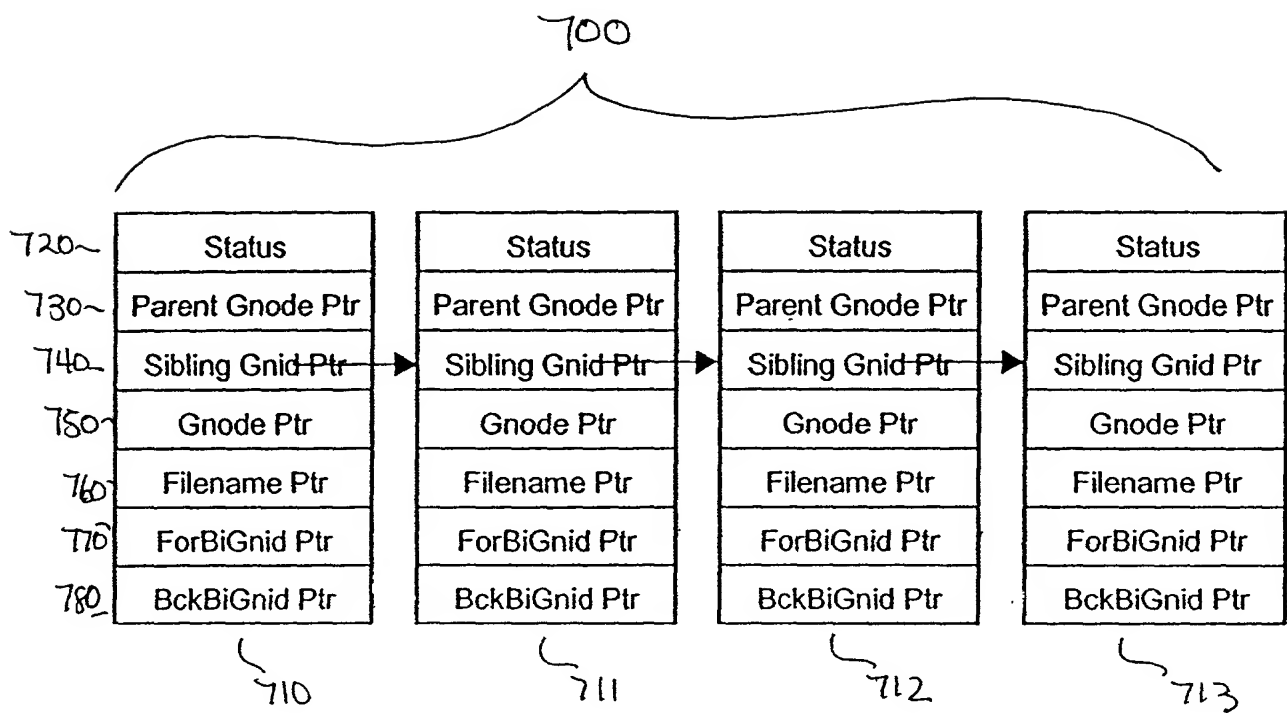


FIGURE 7- Structure of a Gnid String

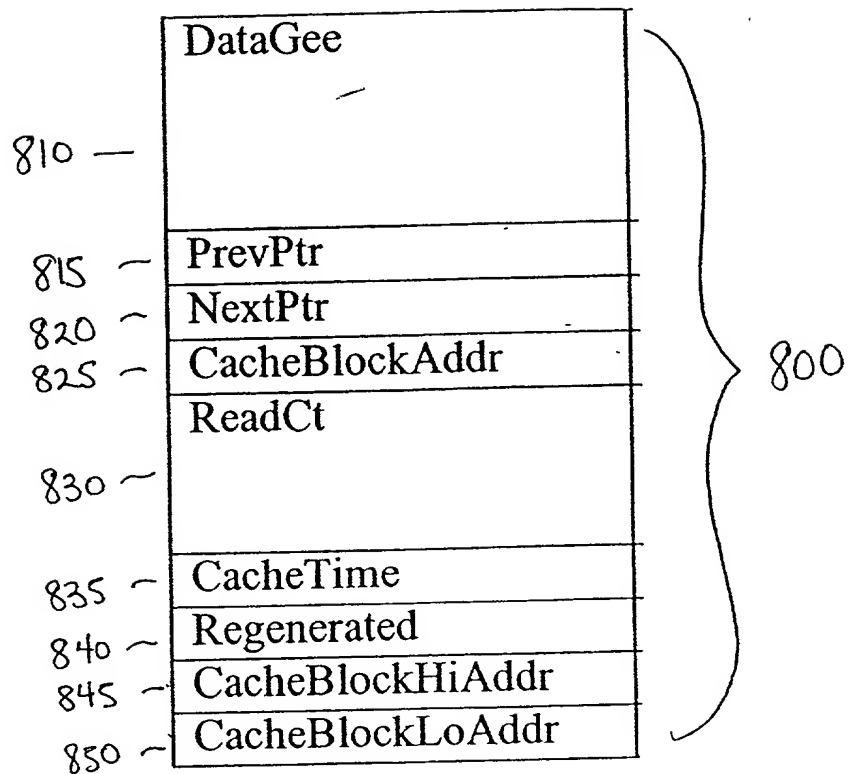


FIGURE 8a - Structure of a Cache Node

350

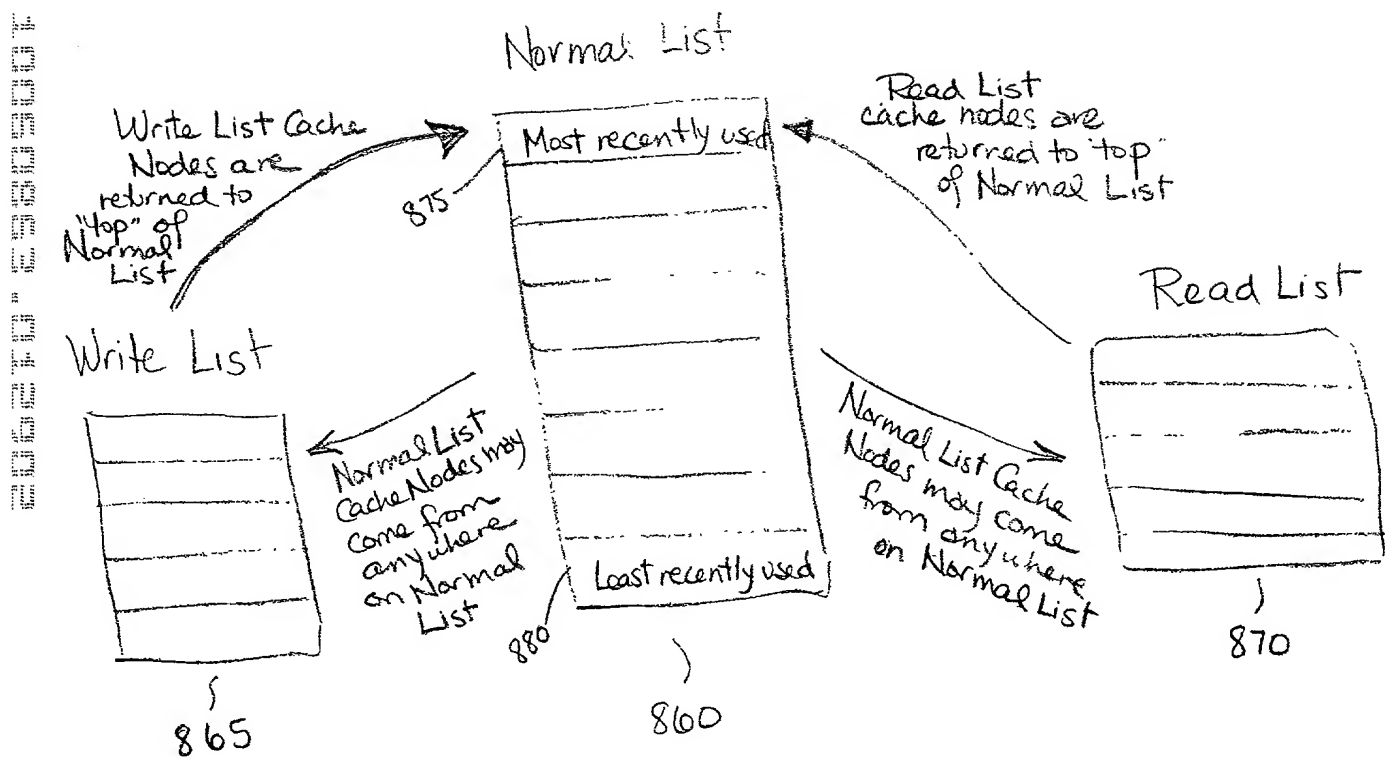


FIGURE 8B - Conceptual division of a Cache Node Table into Three Lists

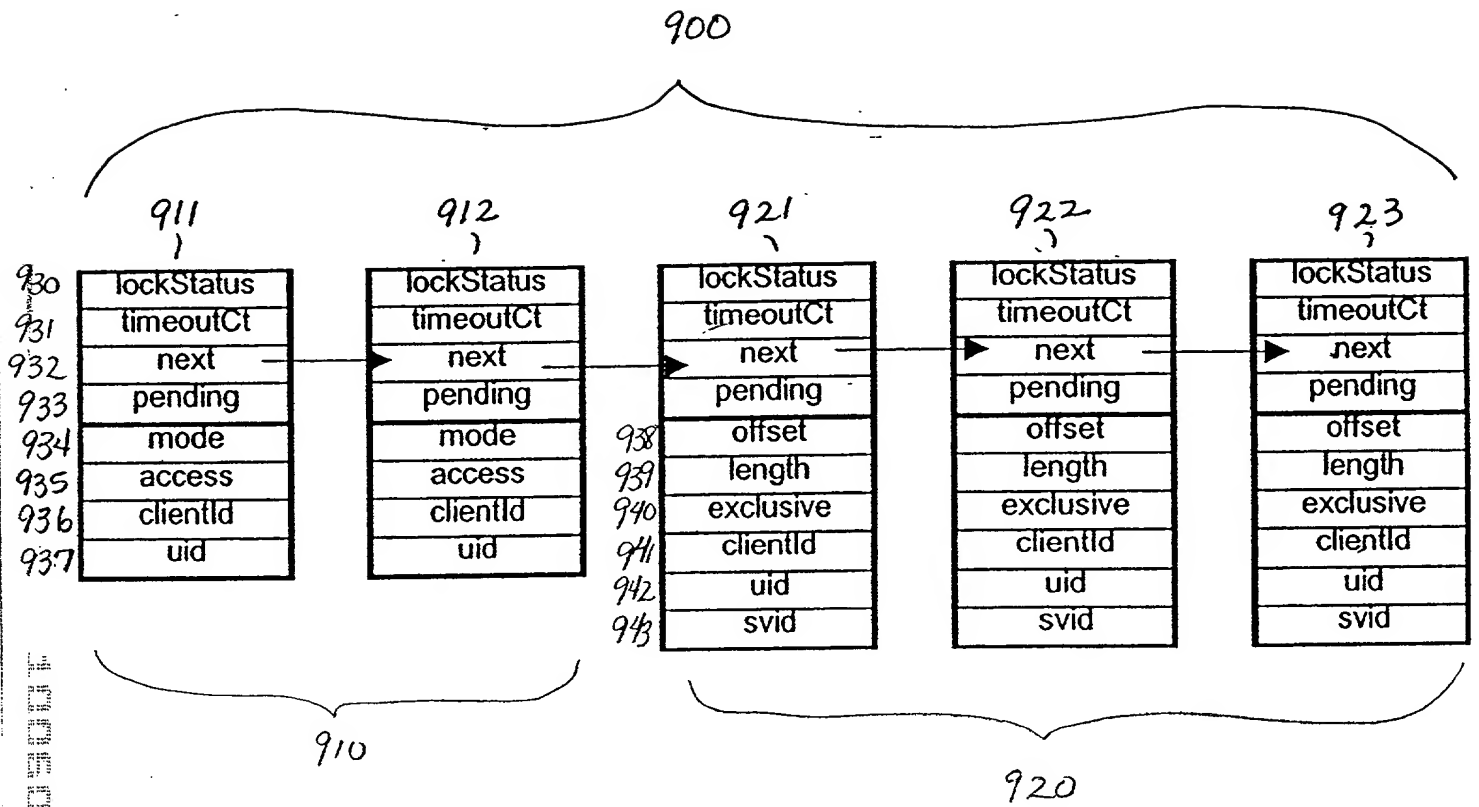


FIGURE 9 - A Sample Lock String

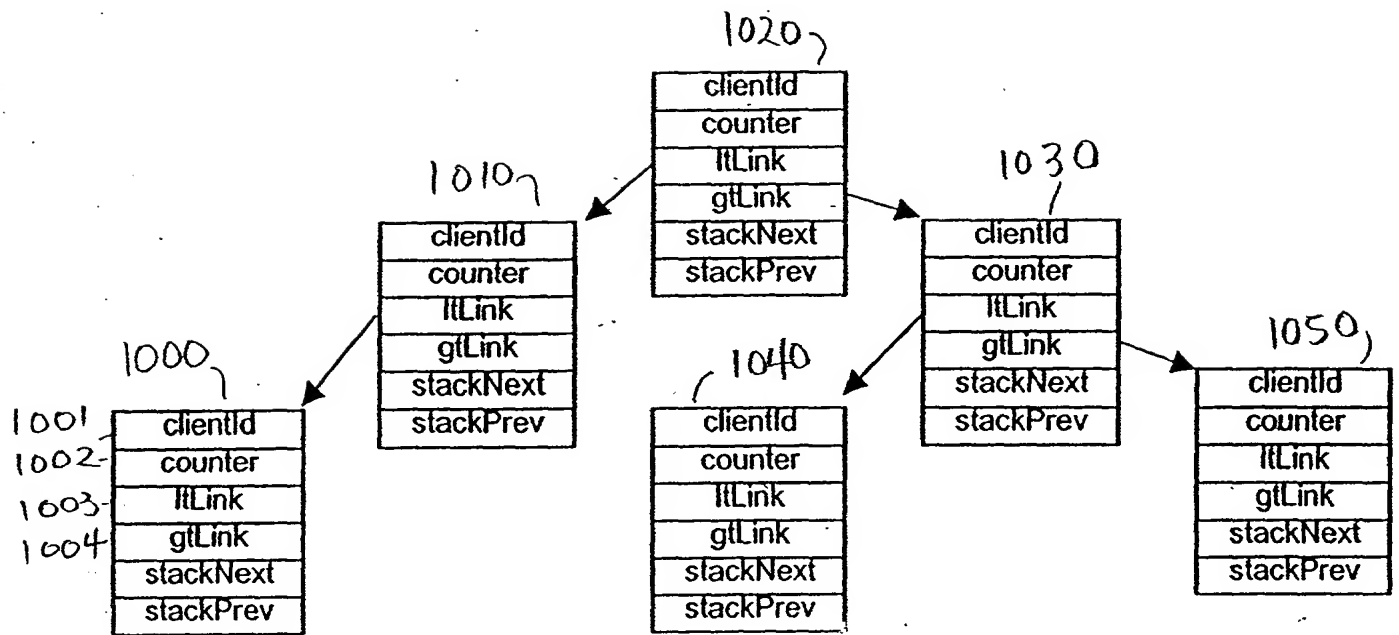


FIGURE 10 - Refresh Nodes configured as a binary tree.

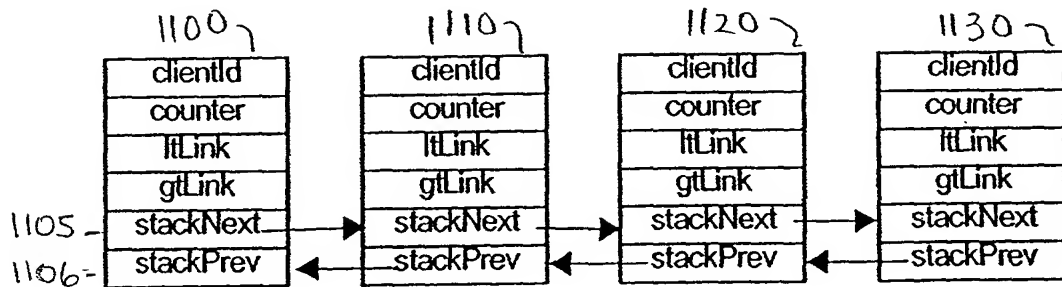


FIGURE 11 - Refresh Nodes configured as a doubly-linked list

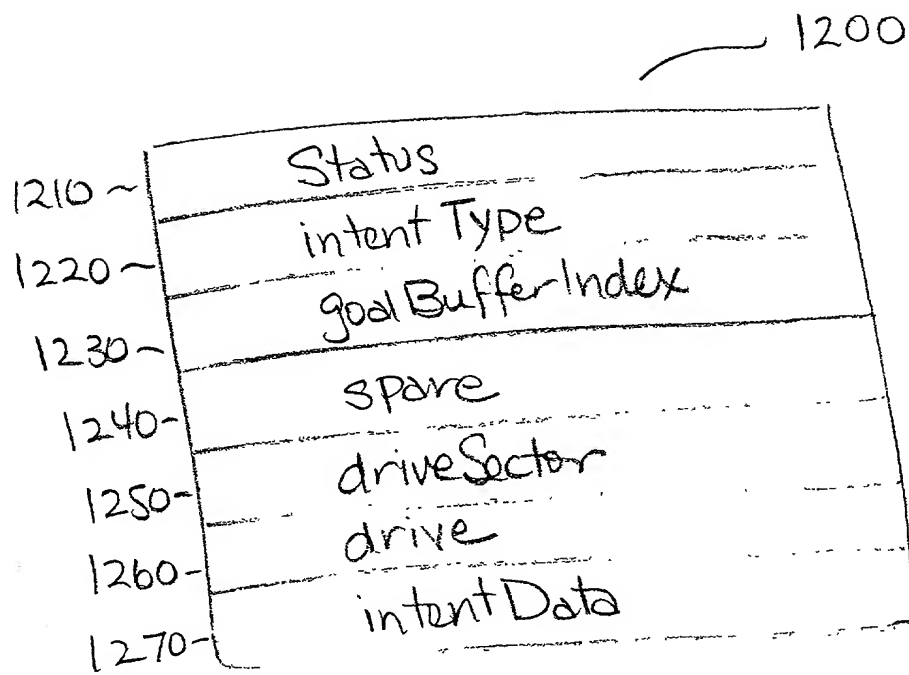


FIGURE 12 - Structure of an Intent Log Entry

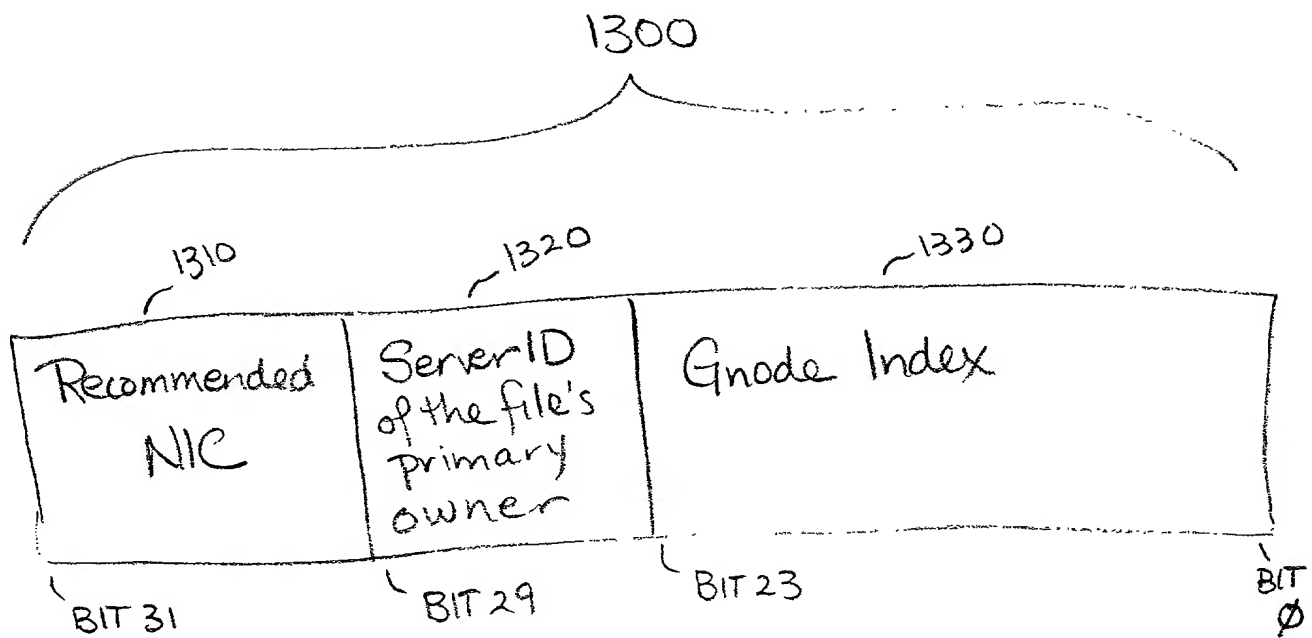


FIGURE 13 - Structure of a File Handle

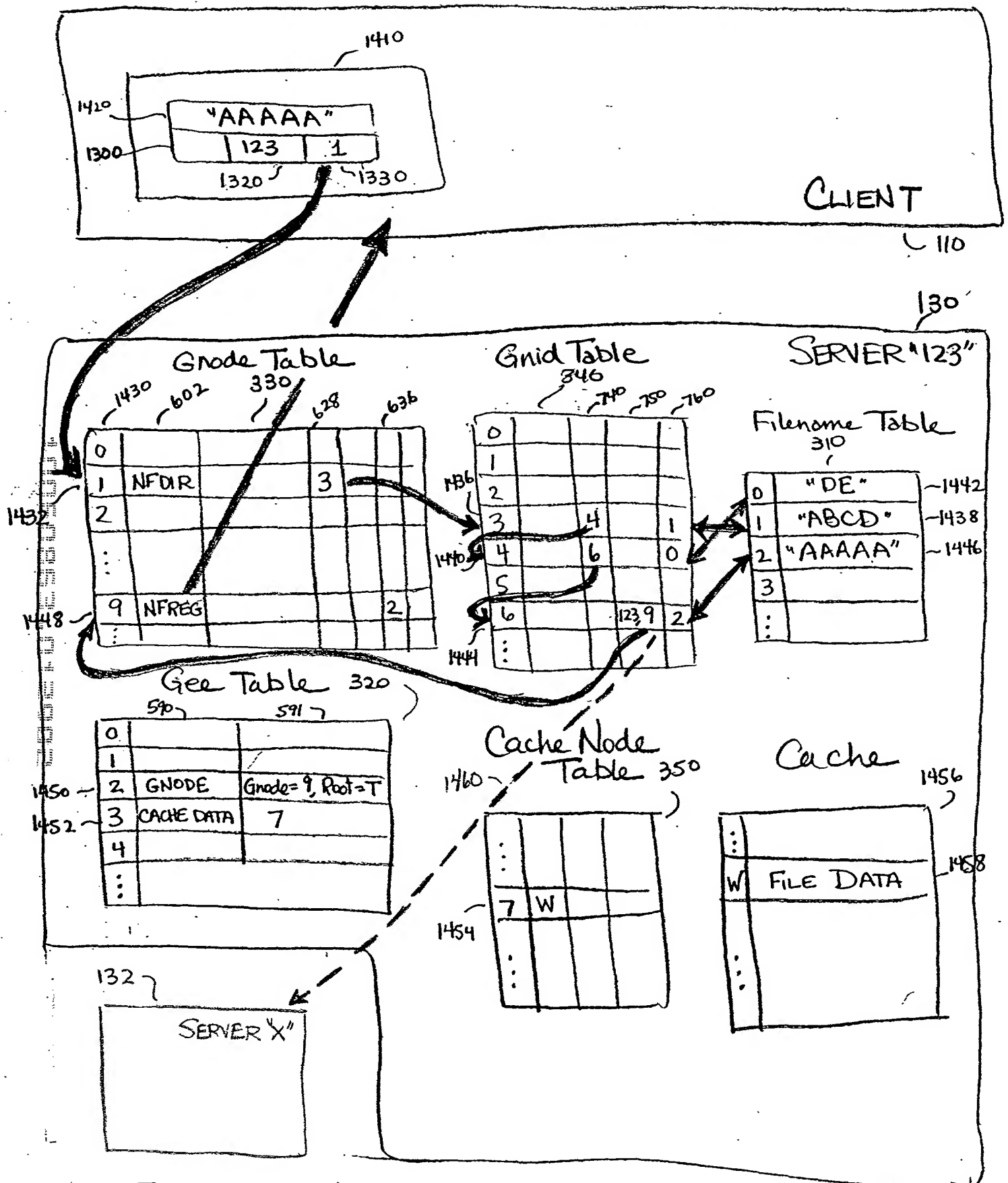


FIGURE 14a: Example of a File Look-Up

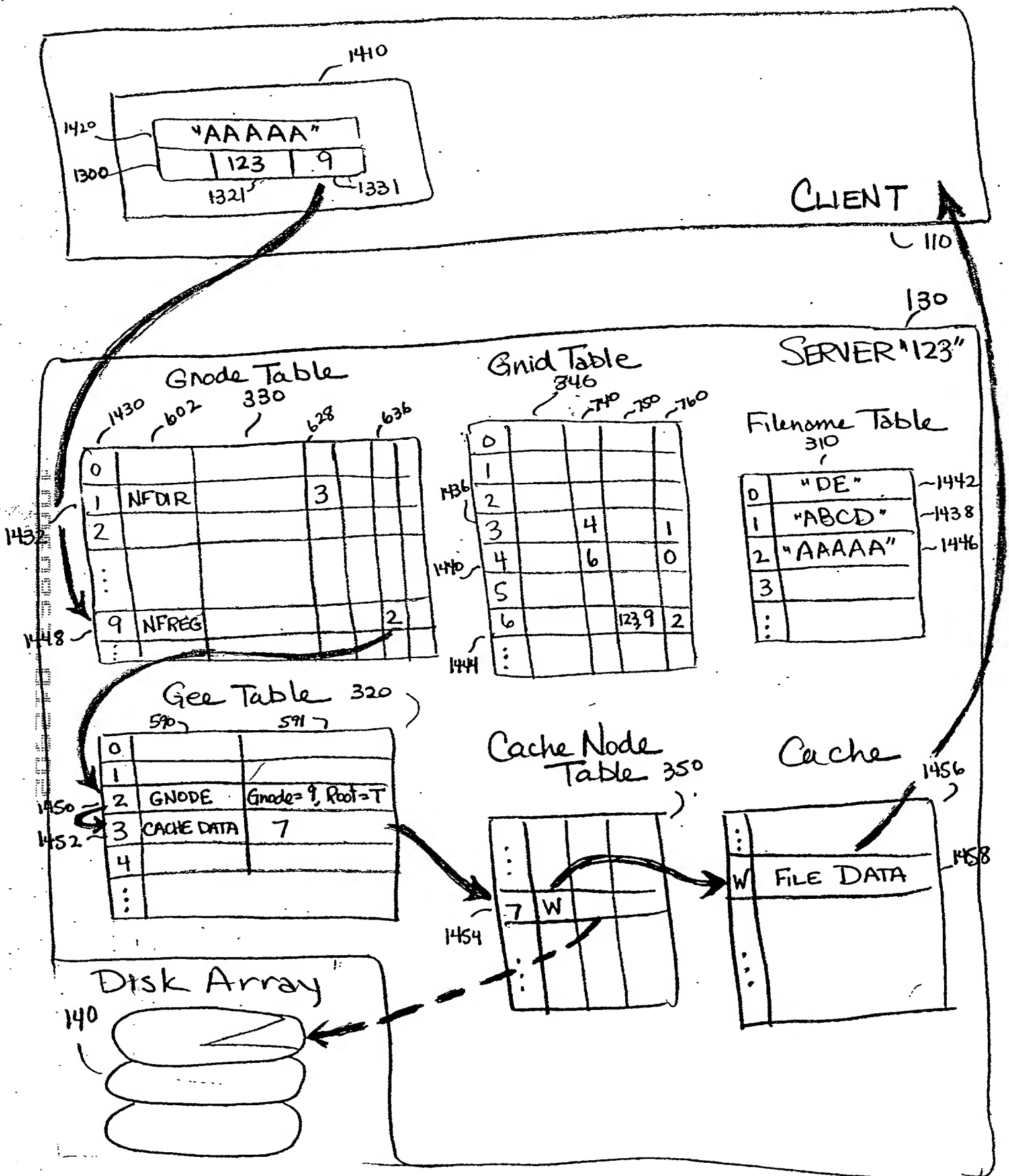


FIGURE 14b Example of a File Access

1500

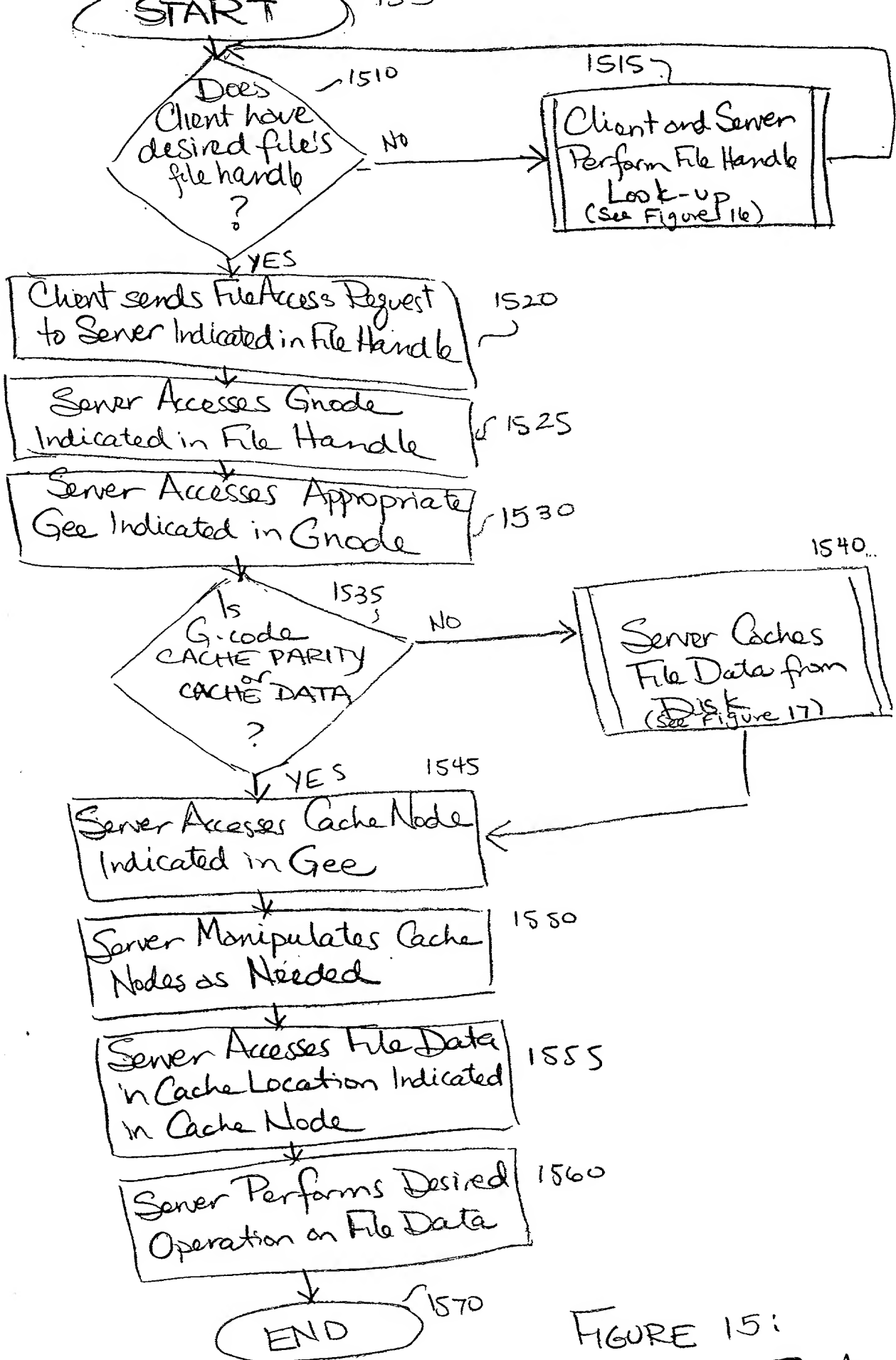


FIGURE 15:
Performing a File Access

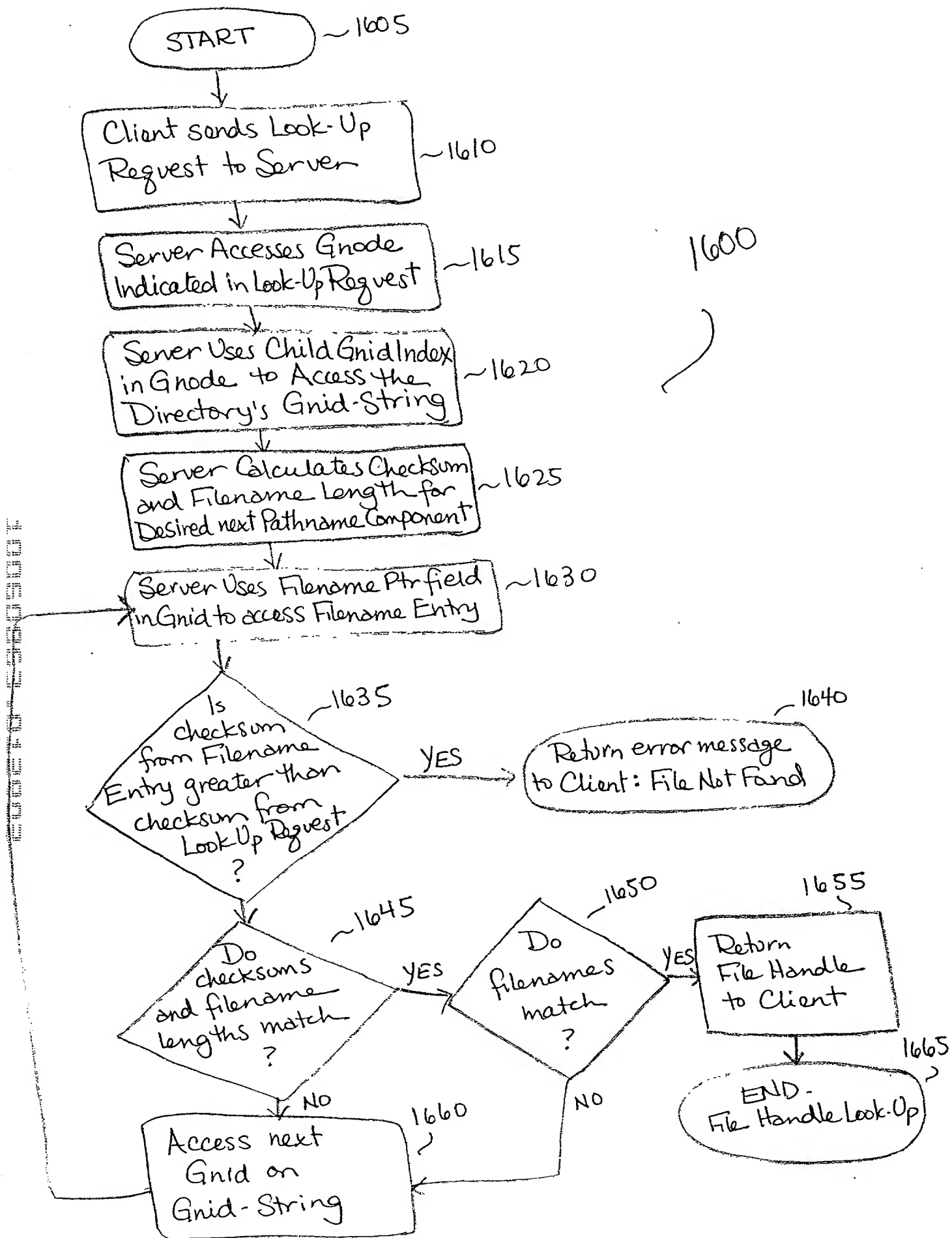


FIGURE 16: Performing a File Handle Look-Up

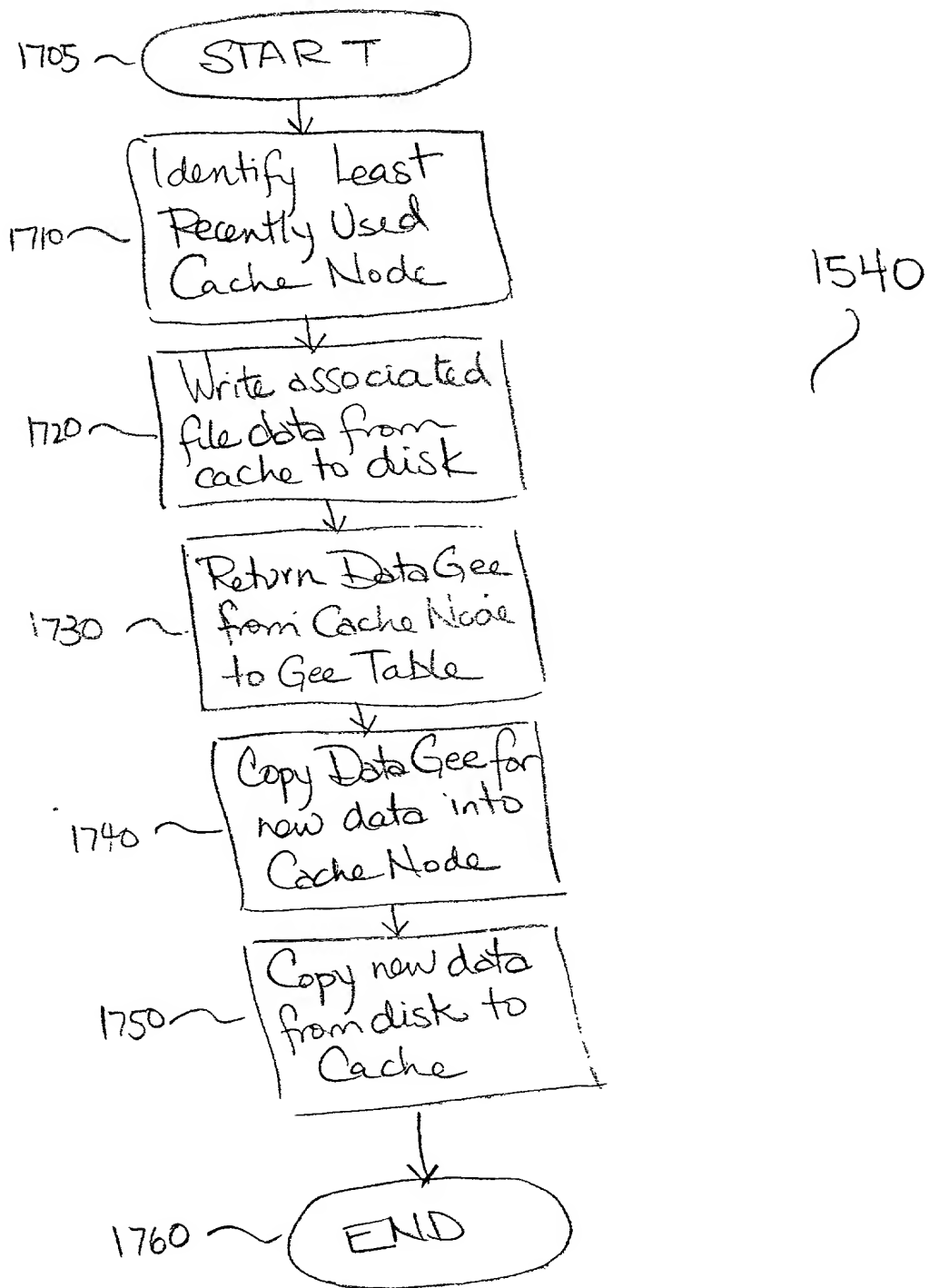


FIGURE 17: Caching File Data

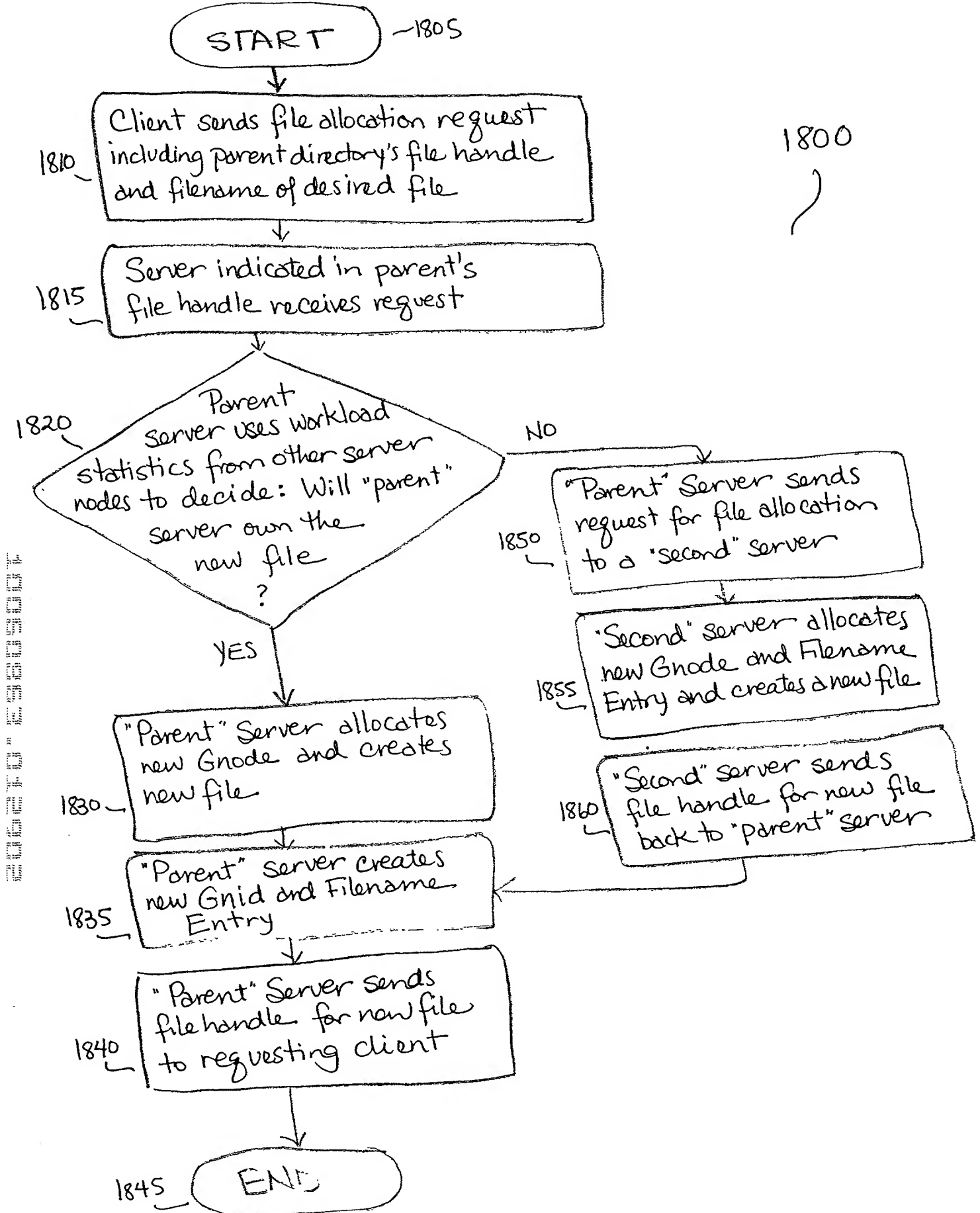


FIGURE 18 - File Allocation

- Gnode Redirectors (GNR)

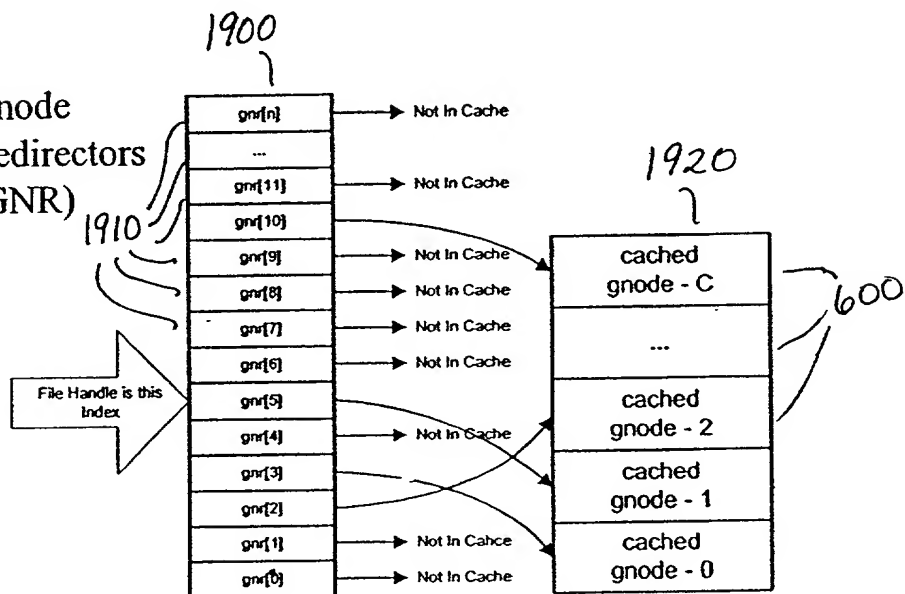


FIGURE 19

2000

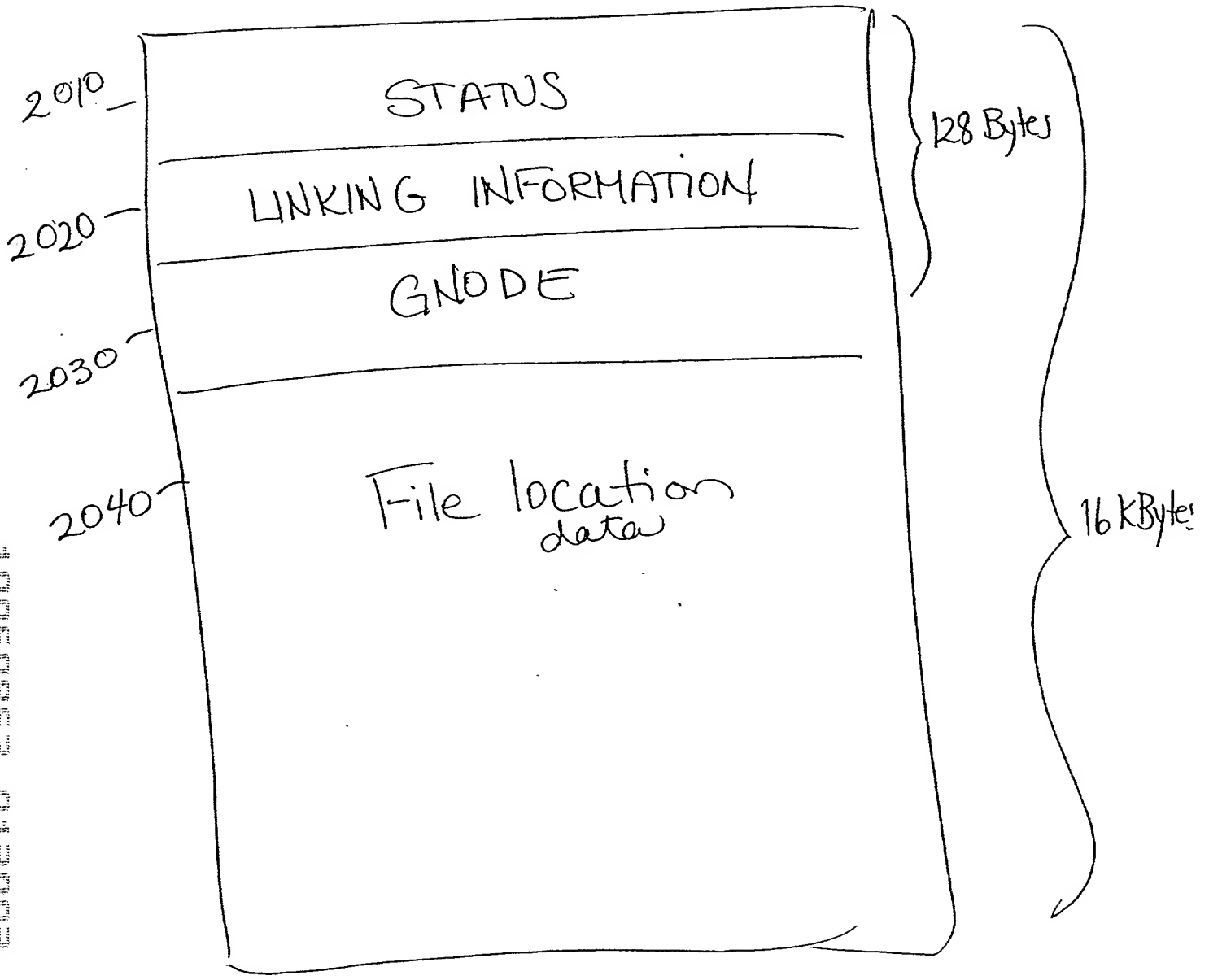


Figure 20a

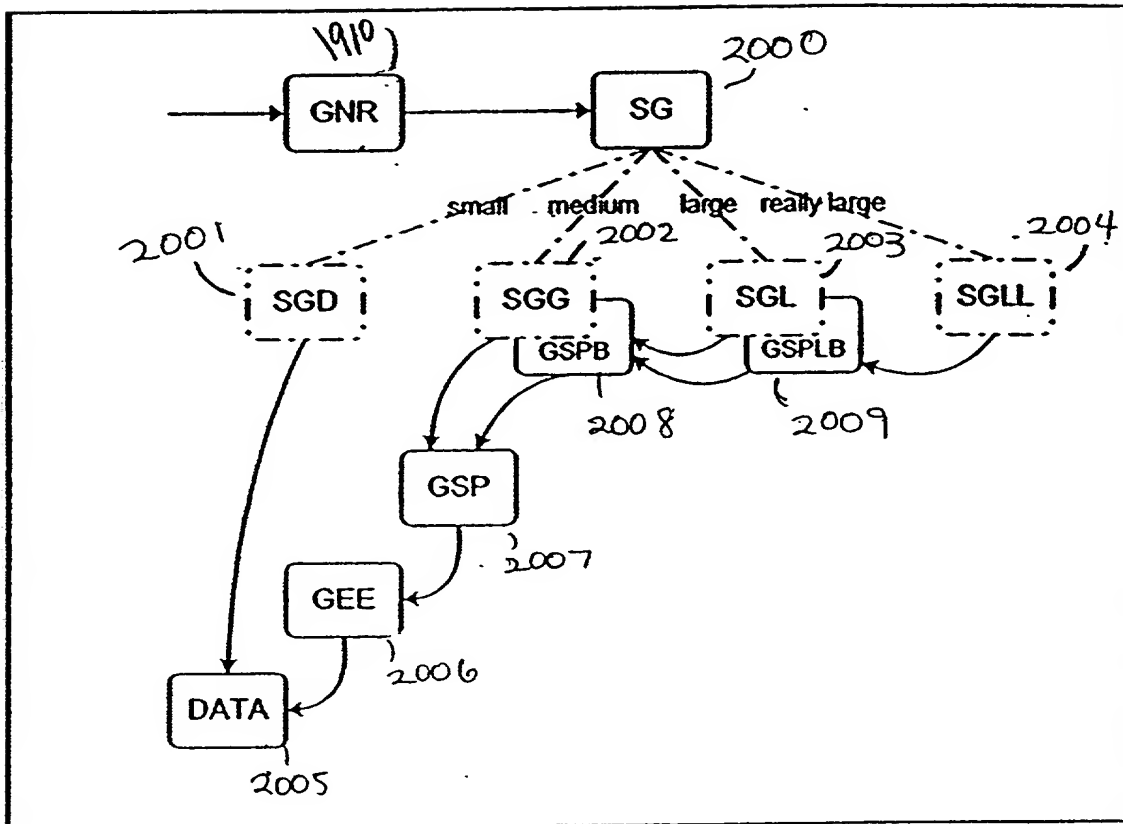


FIGURE 20b

CONVENTIONAL RAID MAPPING
(PRIOR ART)

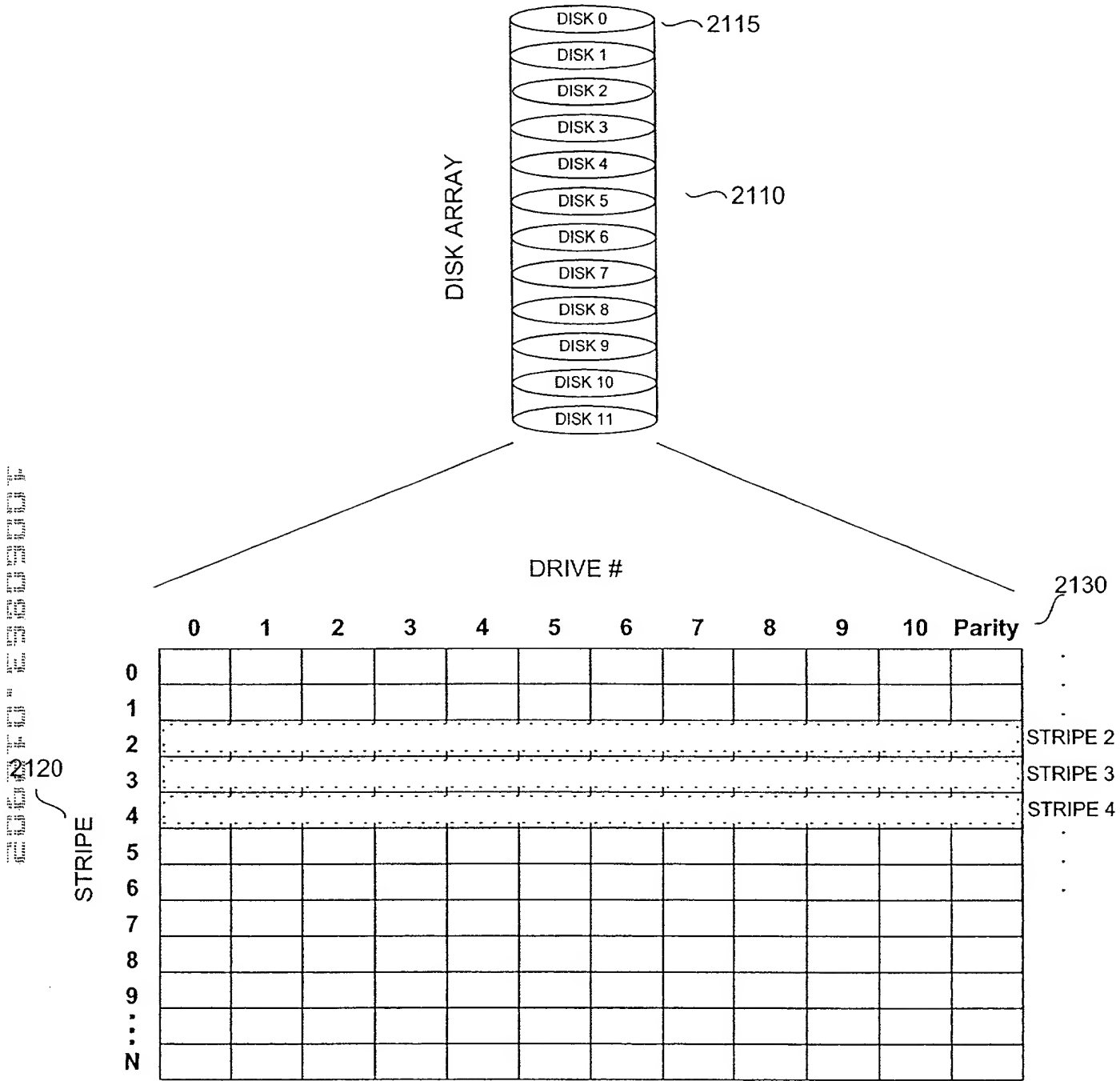


FIGURE 21

FIGURE 22A

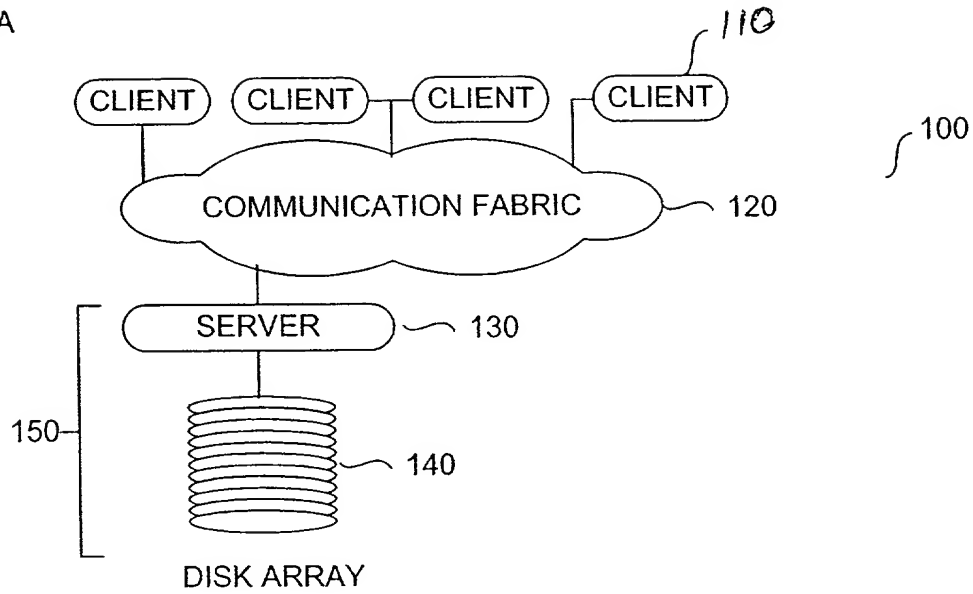
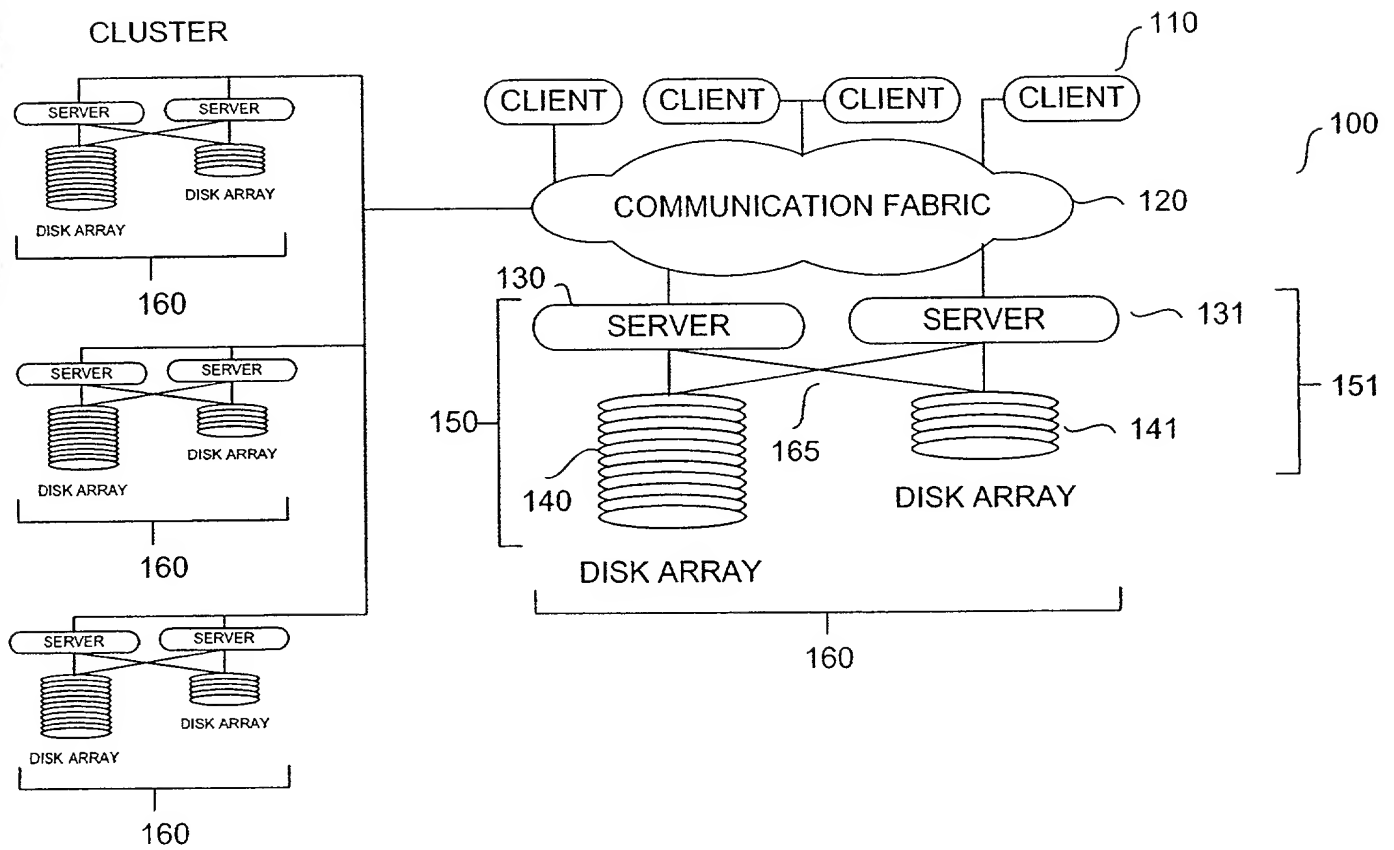


FIGURE 22B



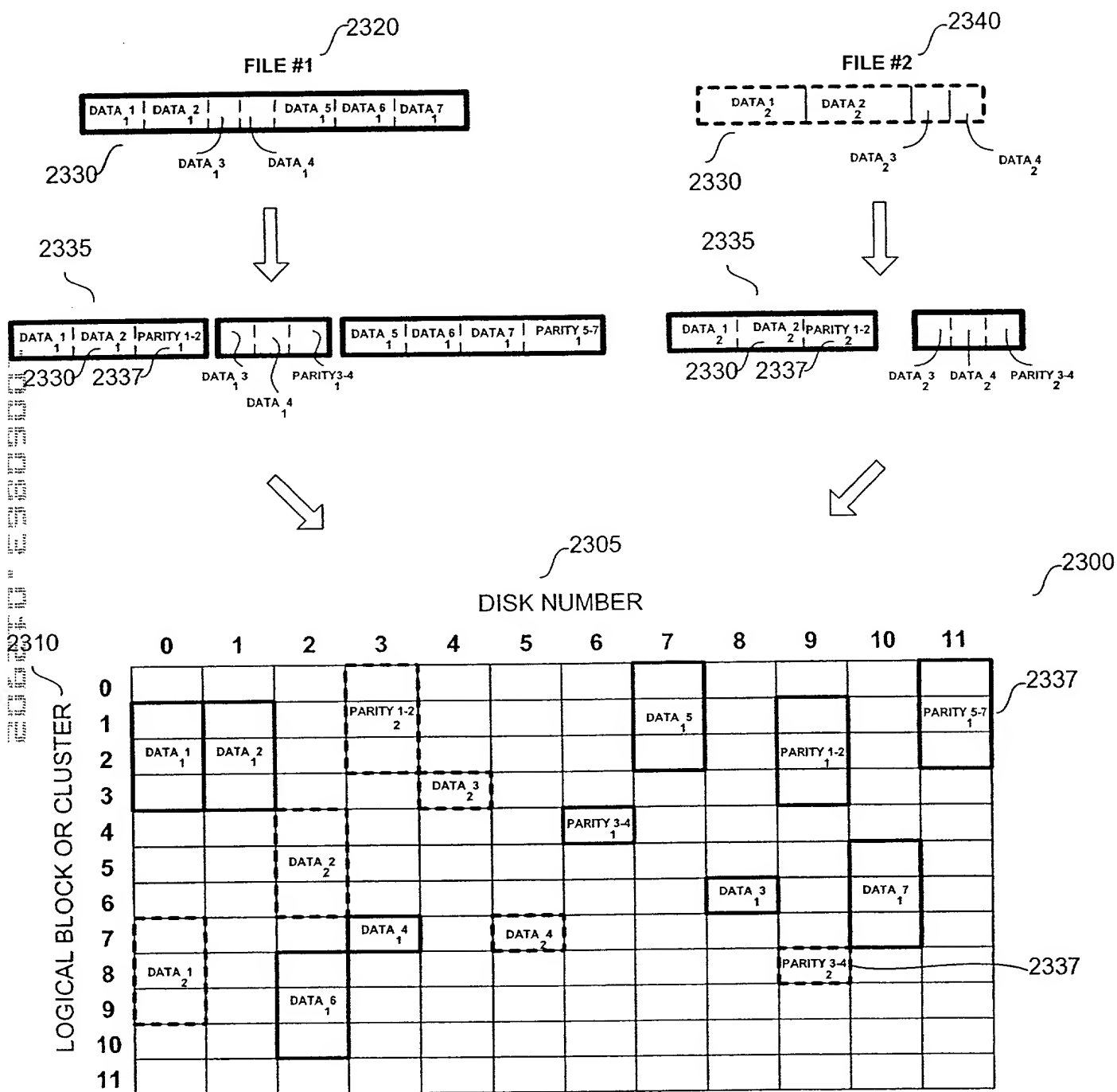
[illegible]

FIGURE 24A

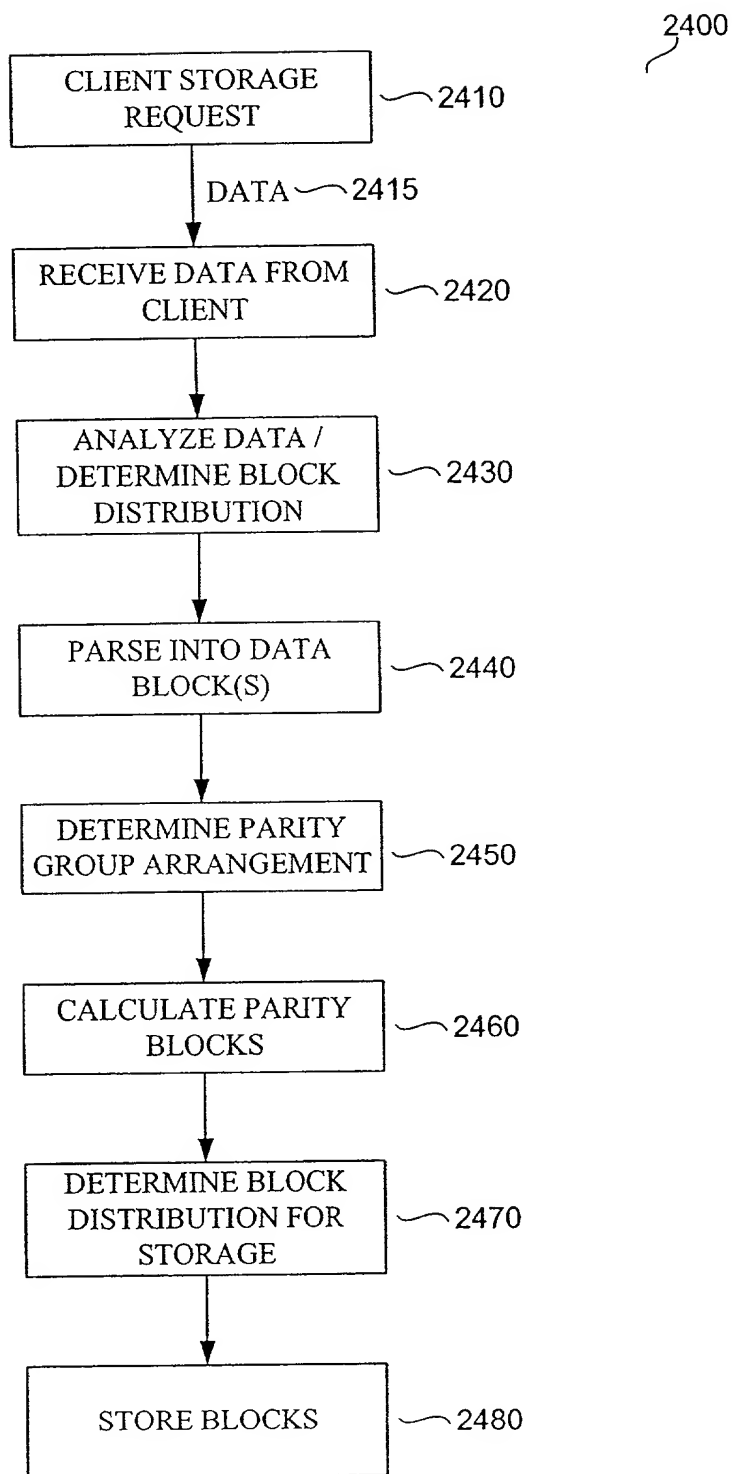


FIGURE 24B

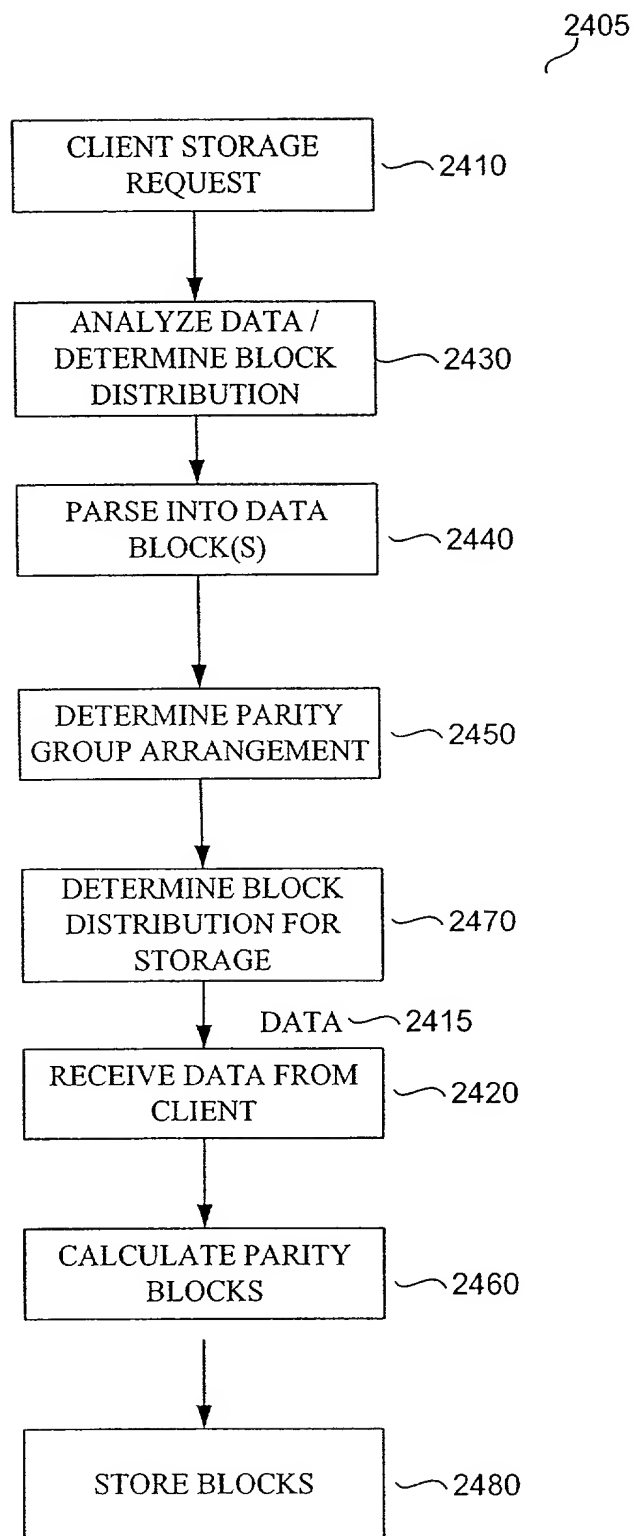


FIGURE 25

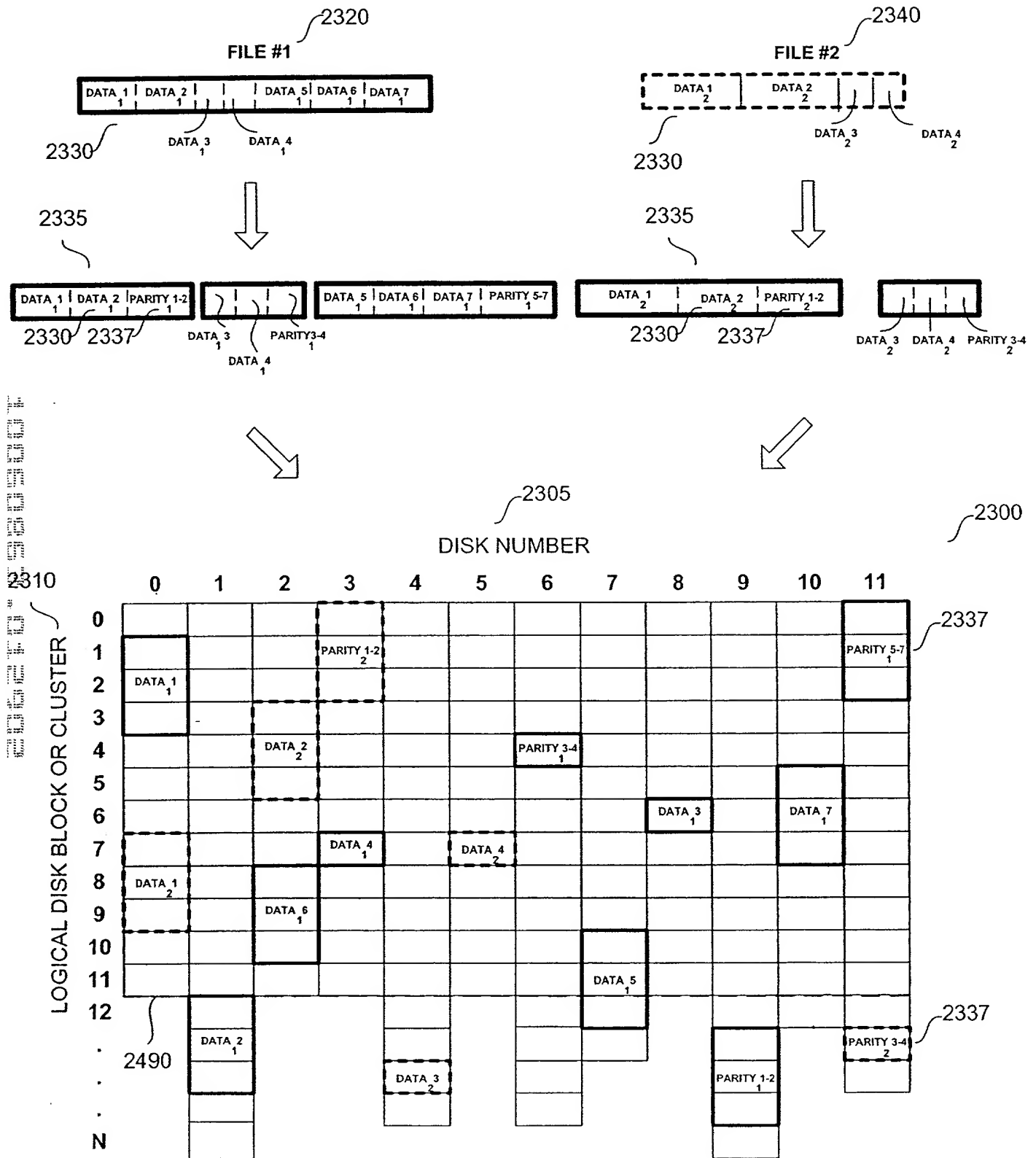


FIGURE 26A

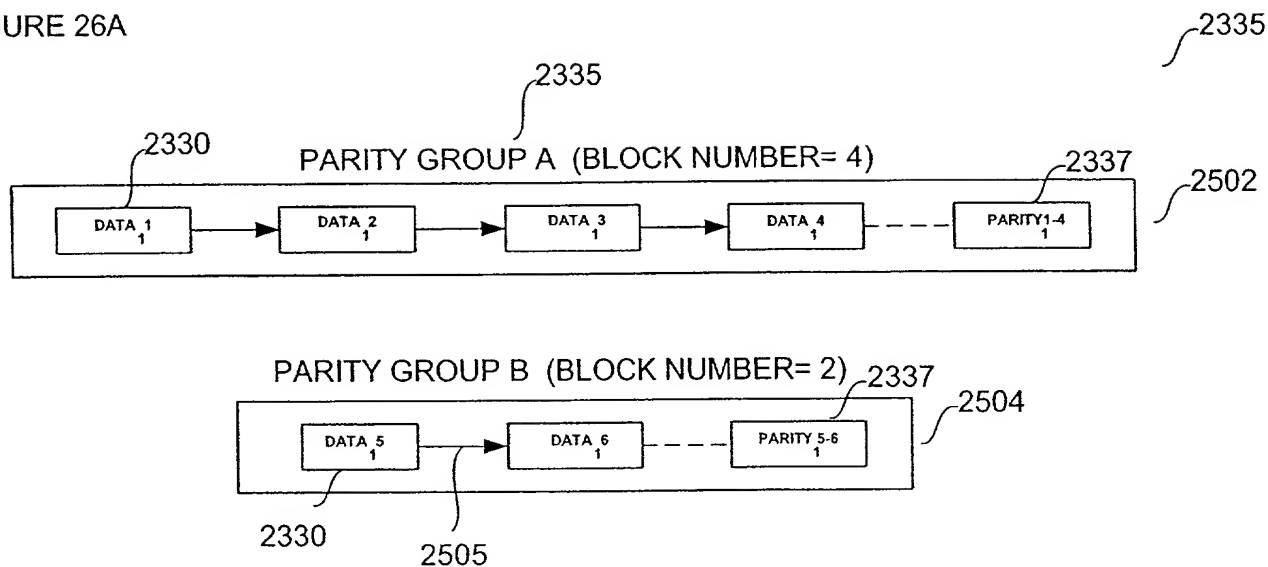
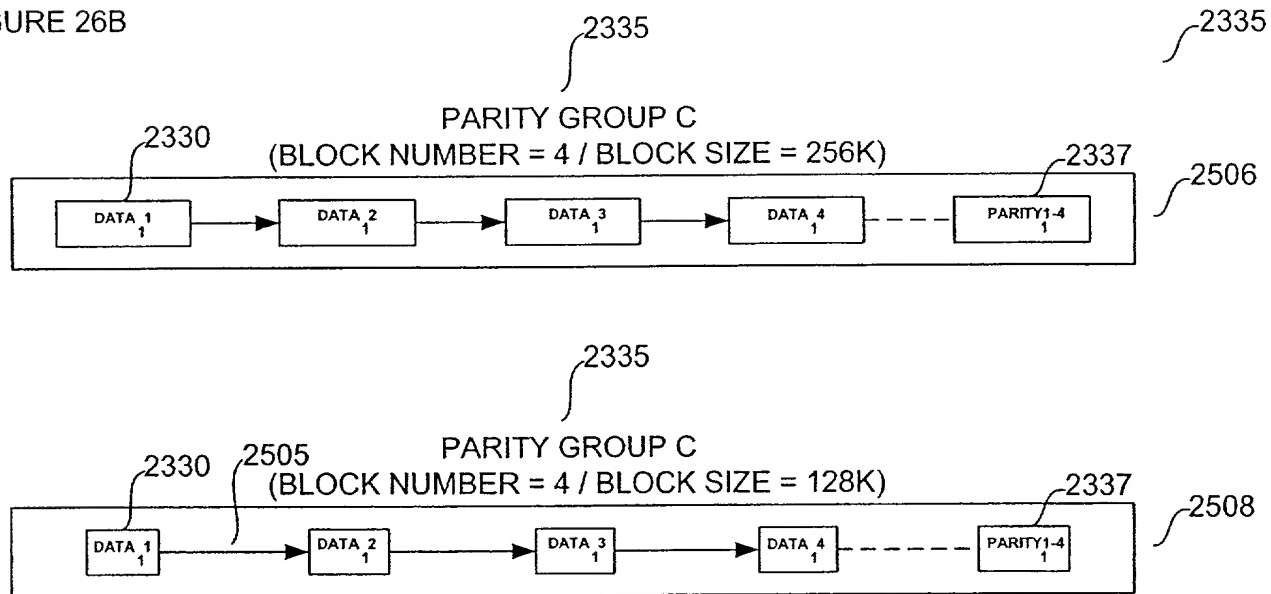


FIGURE 26B



DISK ARRAY INITIALIZATION USING GEE TABLE SPACE ALLOCATION

2530

2532 <u>INDEX</u>	2534 <u>G-CODE</u>	2536 <u>DATA</u>	2542
...	2540
45	GNODE	EXTENT=2	
46	DATA	BLOCKS 456, 457: Drive 13	
47	DATA	BLOCKS 667, 668: Drive 15	
48	DATA	BLOCKS 112, 113: Drive 19	
49	PARITY	BLOCKS 554, 555: Drive 2	
...	2540
76	GNODE	EXTENT=3	
77	DATA	BLOCKS 460, 461, 462: Drive 13	
78	DATA	BLOCKS 671, 672, 673: Drive 15	
79	PARITY	BLOCKS 121, 122, 123: Drive 19	
...	2540
88	GNODE	EXTENT=2	
89	DATA	BLOCKS 463, 464, 465: Drive 2	
90	DATA	BLOCKS 674, 675, 676: Drive 5	
91	PARITY	BLOCKS 124, 125, 126: Drive 13	
...	

FIGURE 27

ARRAY PREPARATION / G-TABLE FORMATTING

2448

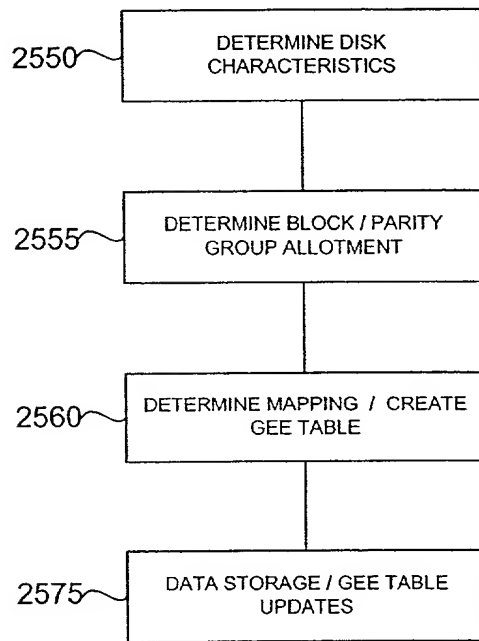


FIGURE 28

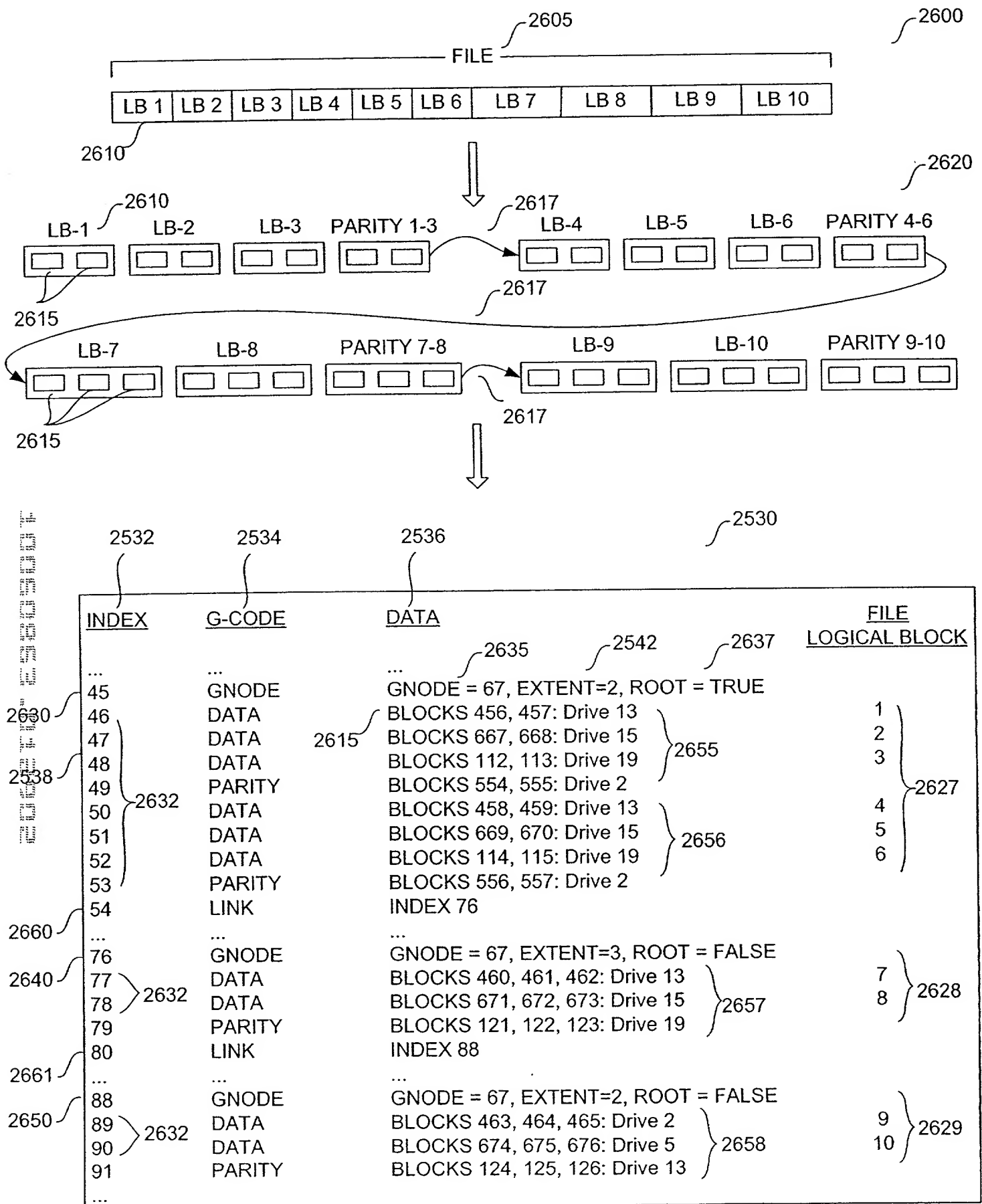


FIGURE 29

DRIVE FAILURE RECOVERY MECHANISM

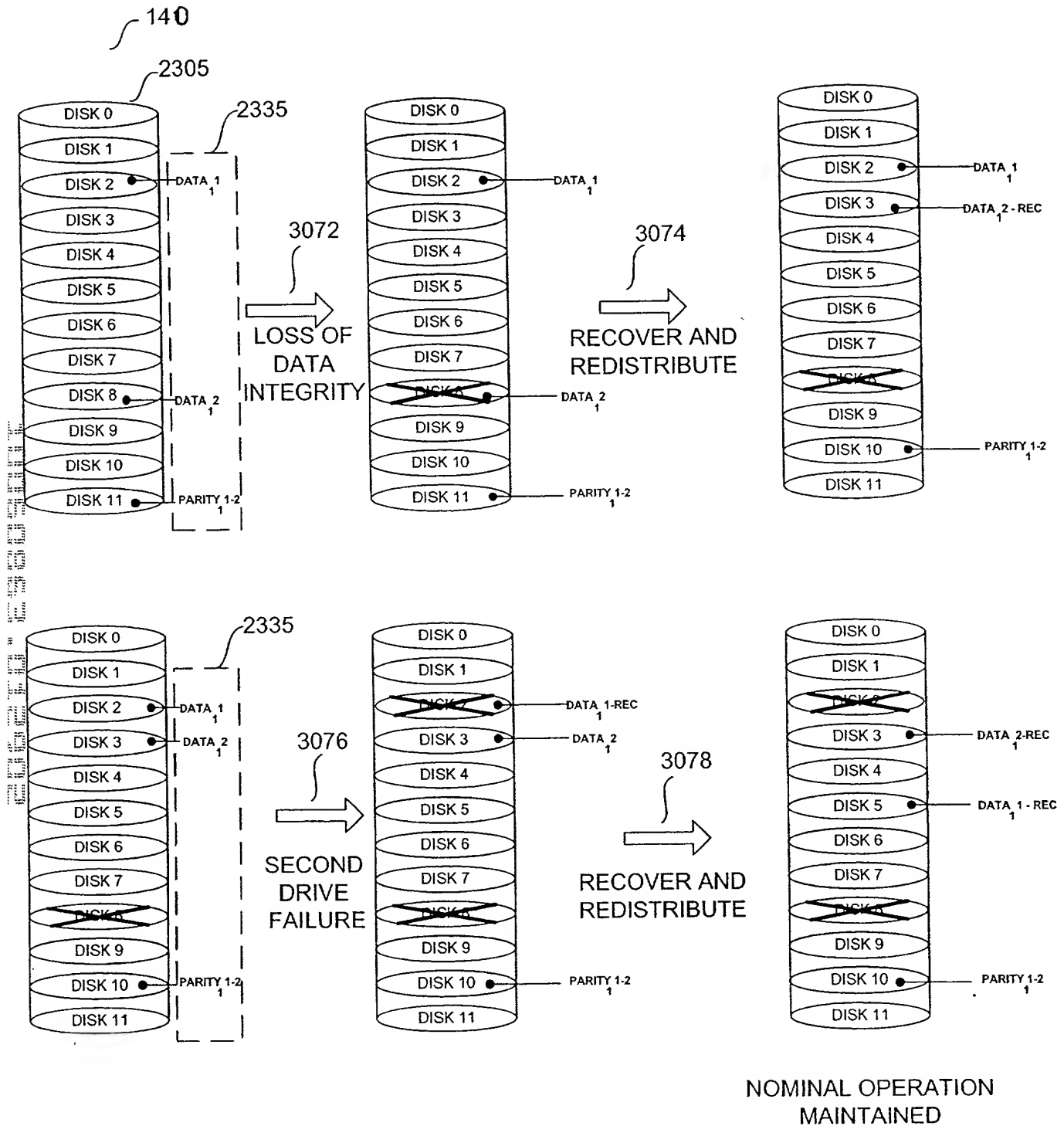


FIGURE 30

DATA RECOVERY
PROCESS 3172

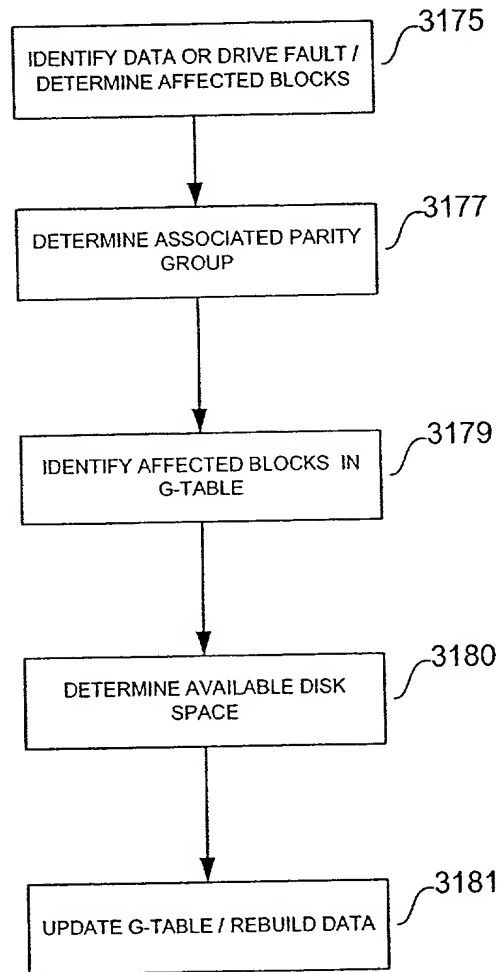
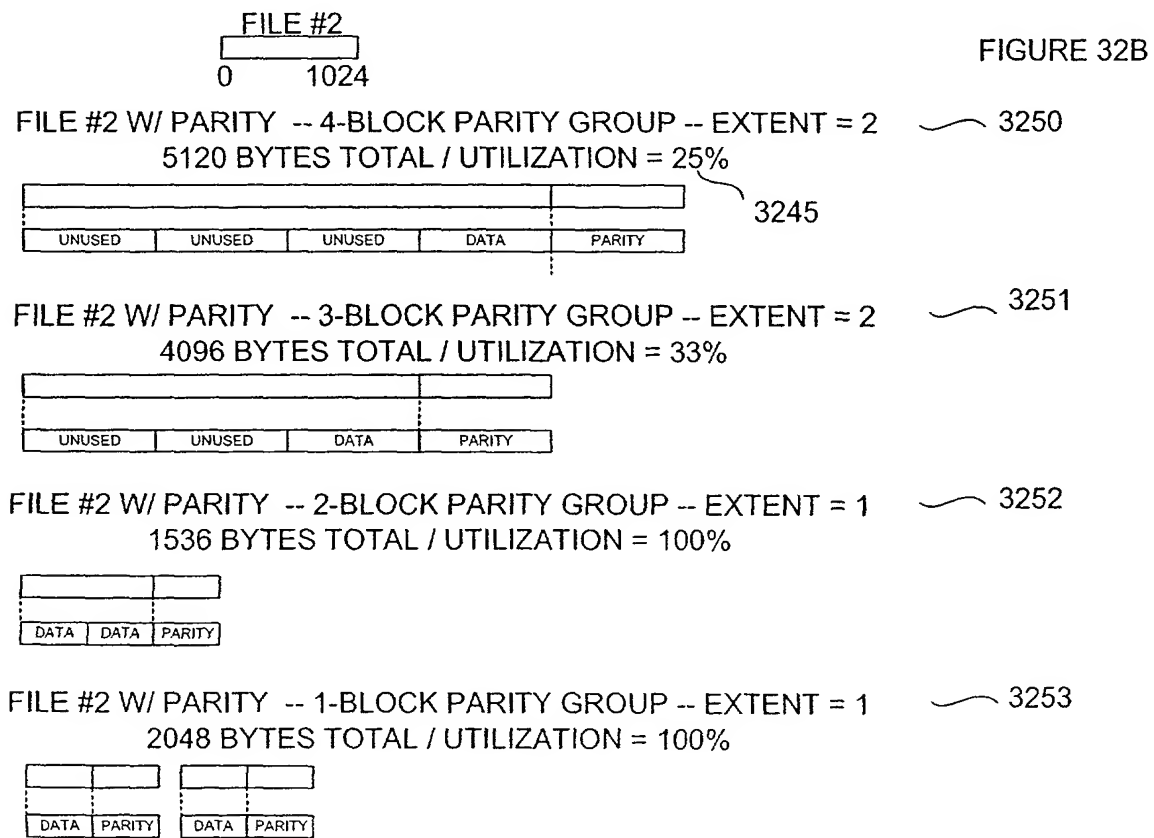
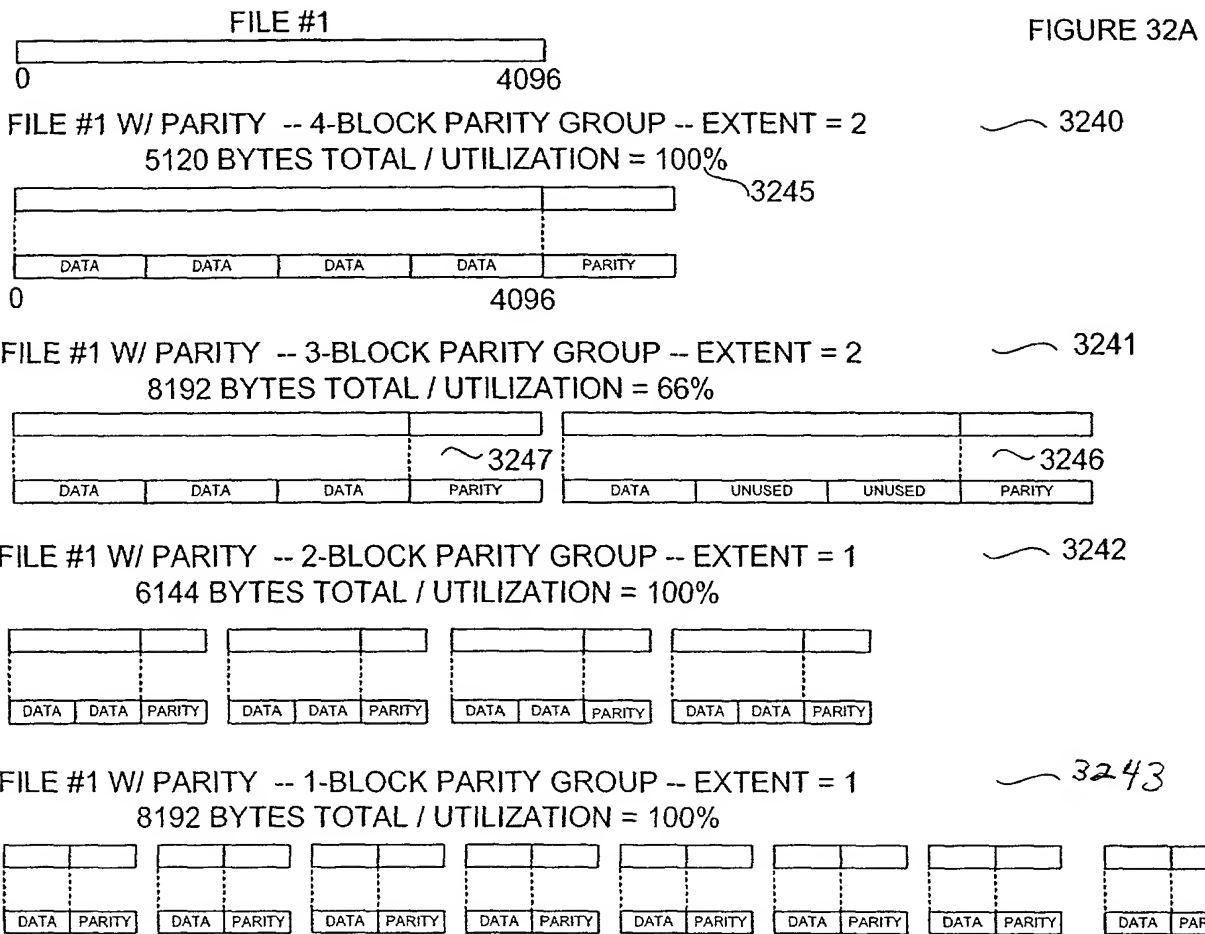


FIGURE 31



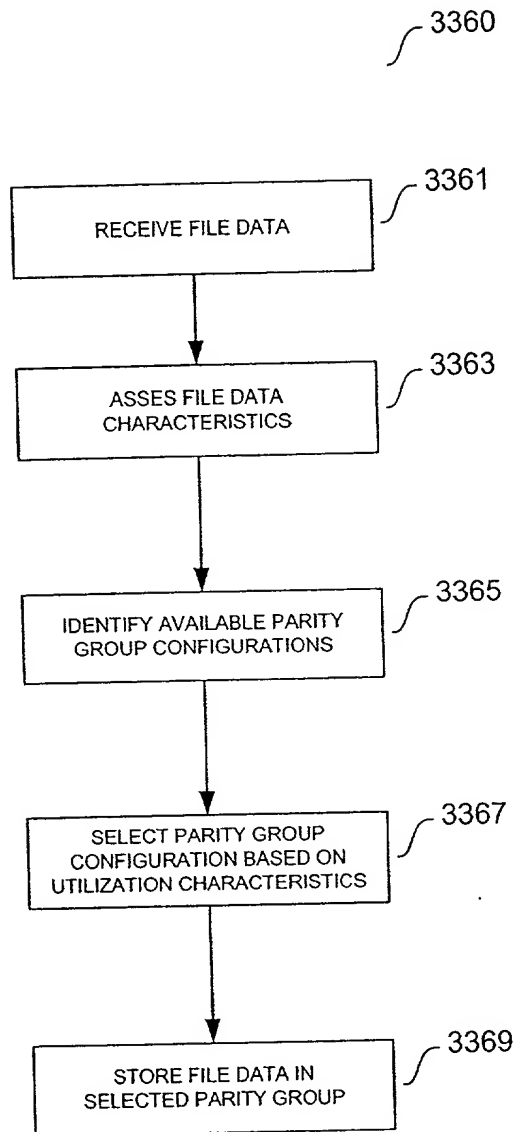


FIGURE 33

FIGURE 34A

INITIAL ALLOCATION				DISK SPACE %
<div>DATA DATA DATA DATA PARITY</div>	4 block parity	10000 groups		36%
<div>DATA DATA DATA PARITY</div>	3 block parity	10000 groups		28%
<div>DATA DATA PARITY</div>	2 block parity	10000 groups		22%
<div>DATA PARITY</div>	1 block parity	10000 groups		14%

FIGURE 34B

DISK USAGE				DISK SPACE %
FREE	OCCUPIED	TOTAL		
2500 groups	7500 groups	10000 groups		36%
7500 groups	2500 groups	10000 groups		28%
3500 groups	6500 groups	10000 groups		22%
500 groups	9500 groups	10000 groups		14%

FIGURE 34C

REDISTRIBUTION				DISK SPACE %
FREE	OCCUPIED	TOTAL		
2500 groups	7500 groups	10000 groups		36%
2500 groups	2500 groups	5000 groups		14%
3500 groups	6500 groups	10000 groups		22%
10500 groups	9500 groups	20000 groups		28%

-5000 groups of 3 block parity

+10000 groups of 1 block parity

REDISTRIBUTION

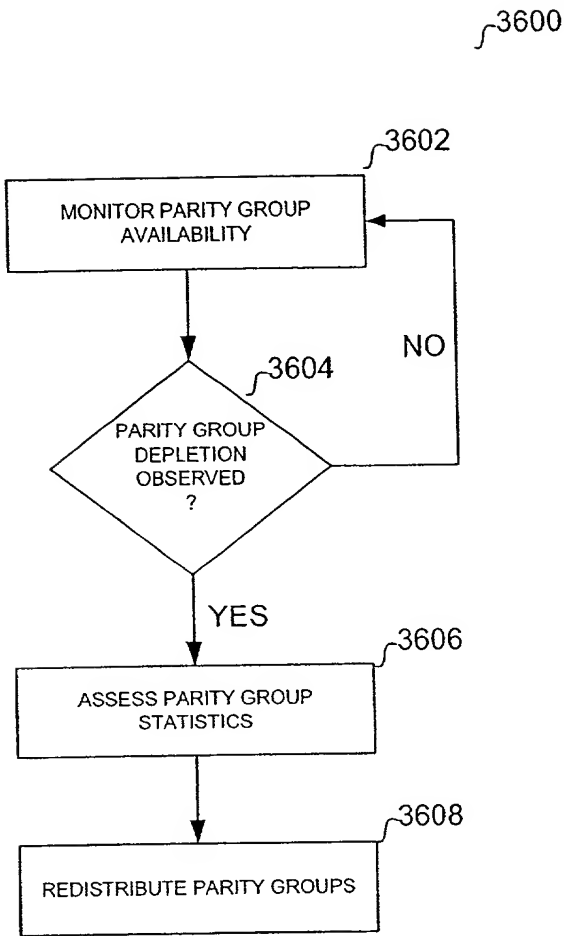


FIGURE 36

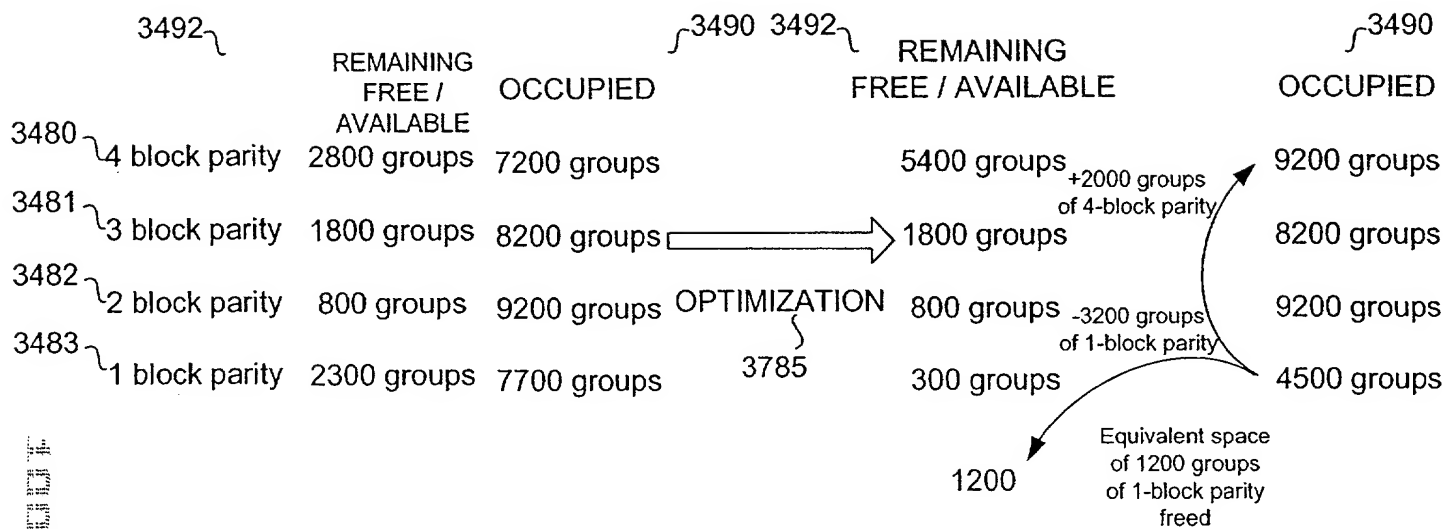


FIGURE 37

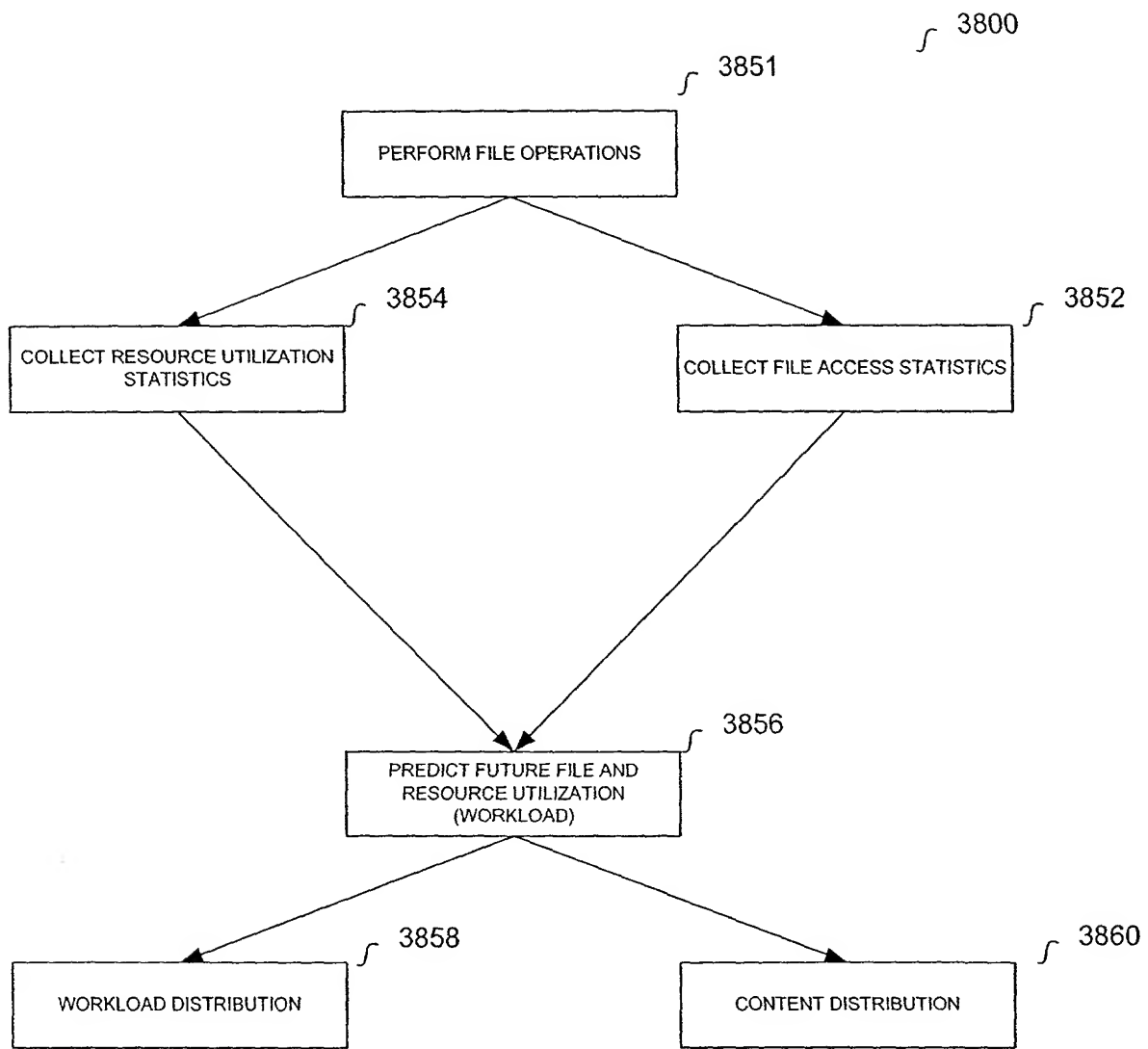


FIGURE 38

3900

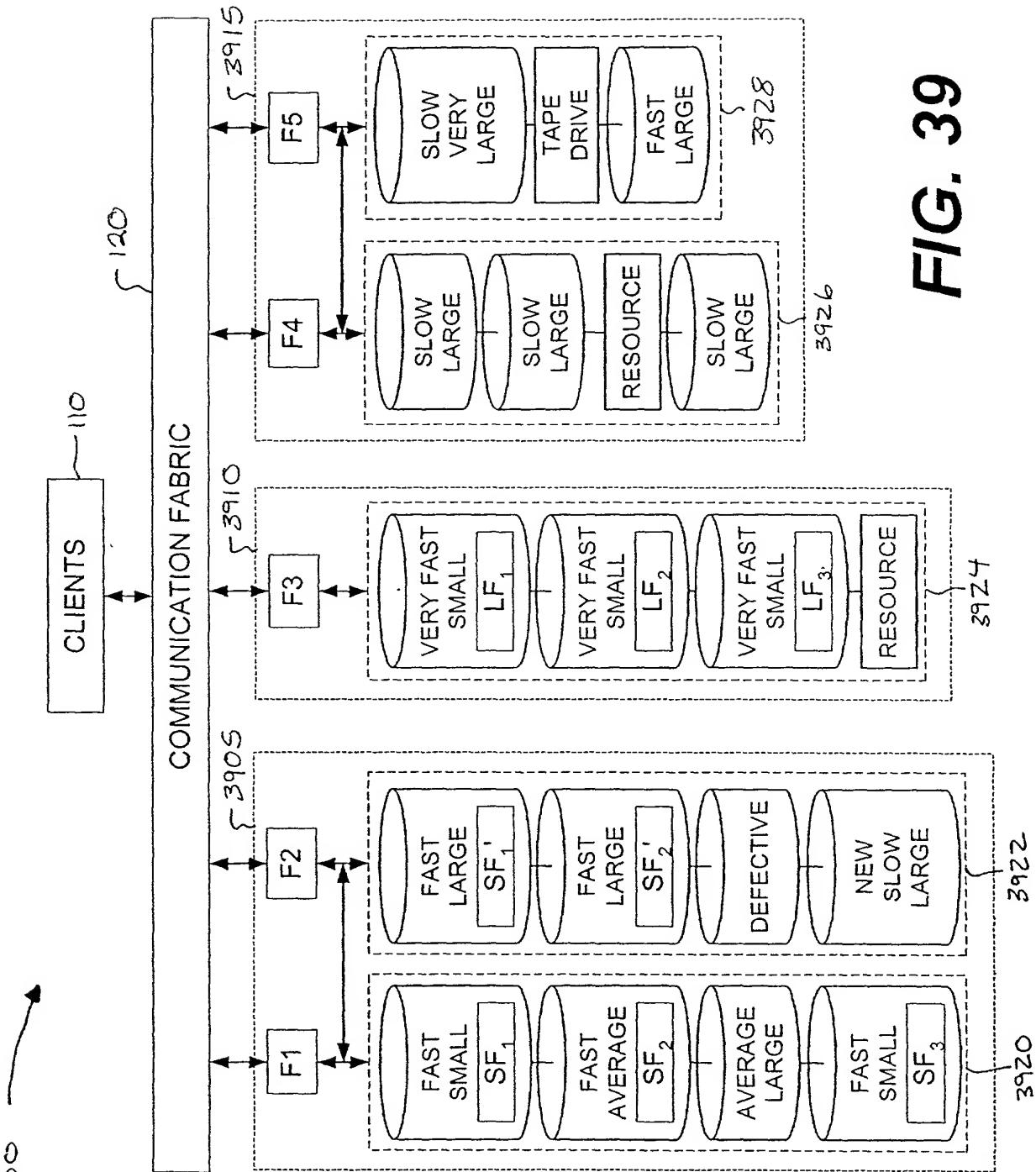


FIG. 39

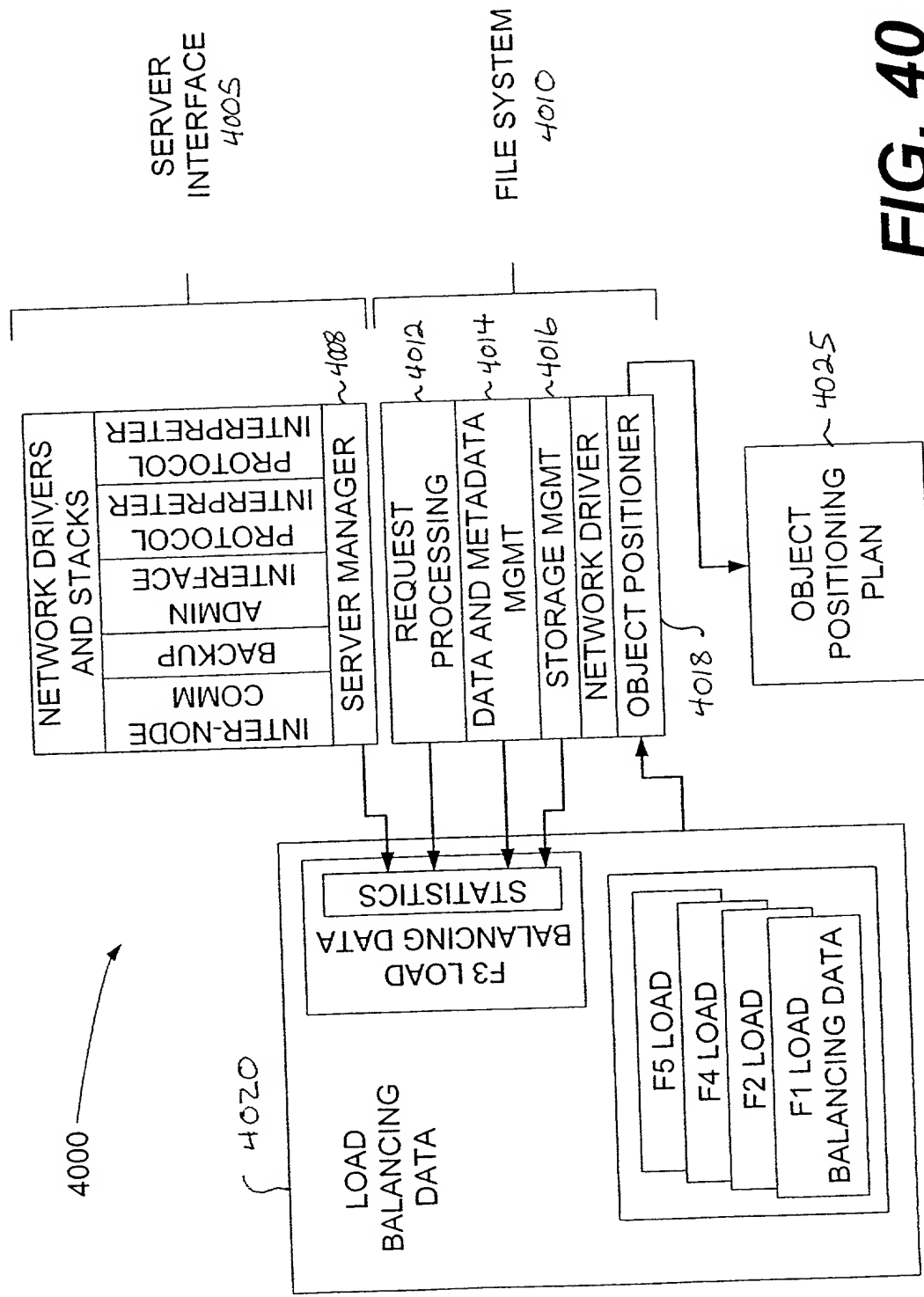


FIG. 40

F3 OBJECT
POSITIONING PLAN

- Push LF to F4-F5 Cluster
- Issue File Handle For LF = Stale
- If Requested,
 - Send acceptance for copy of SF to F1
 - Create copy of SF
 - Send file handle of SF to F1

4025

FIG. 41

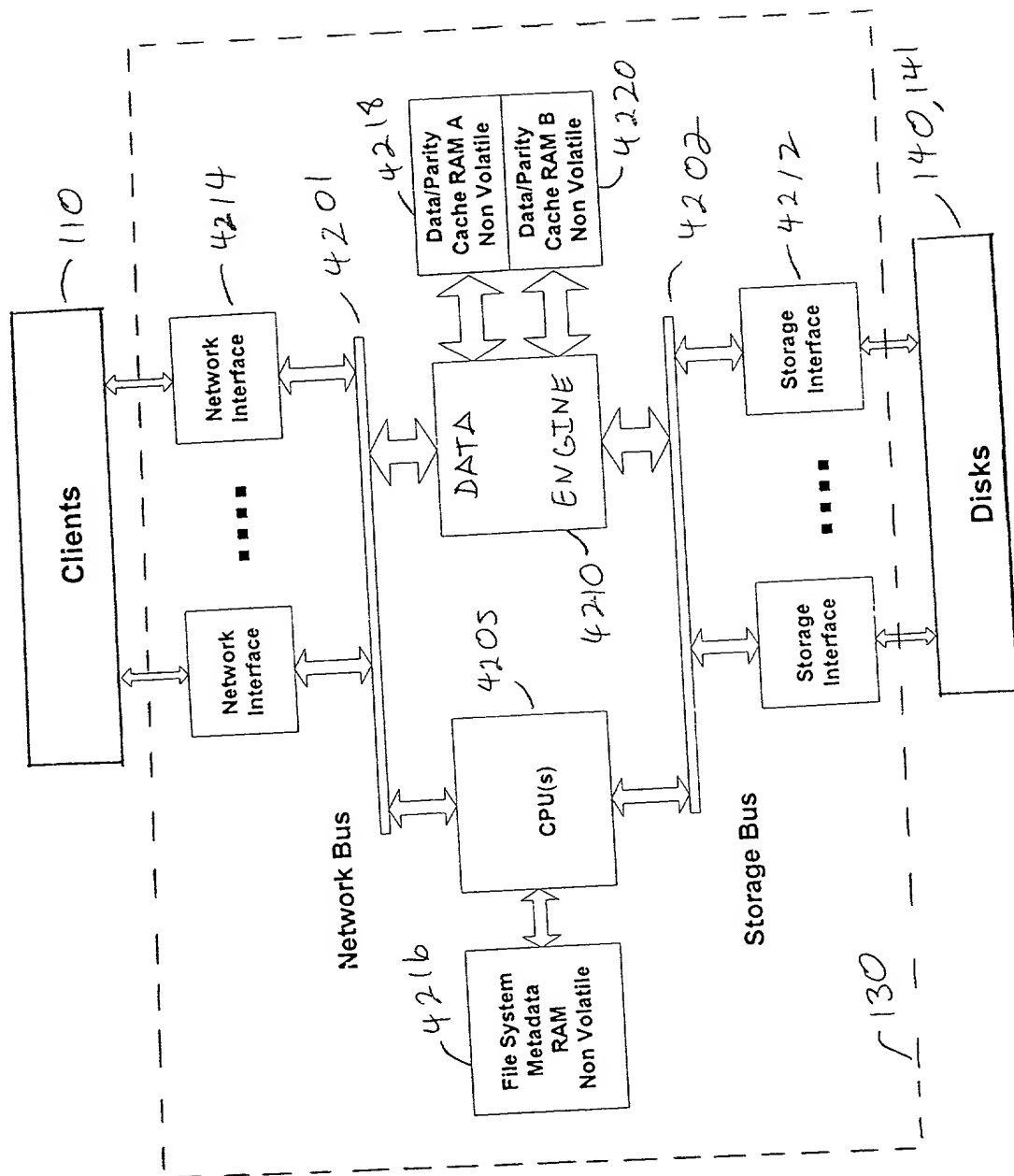


FIGURE 42

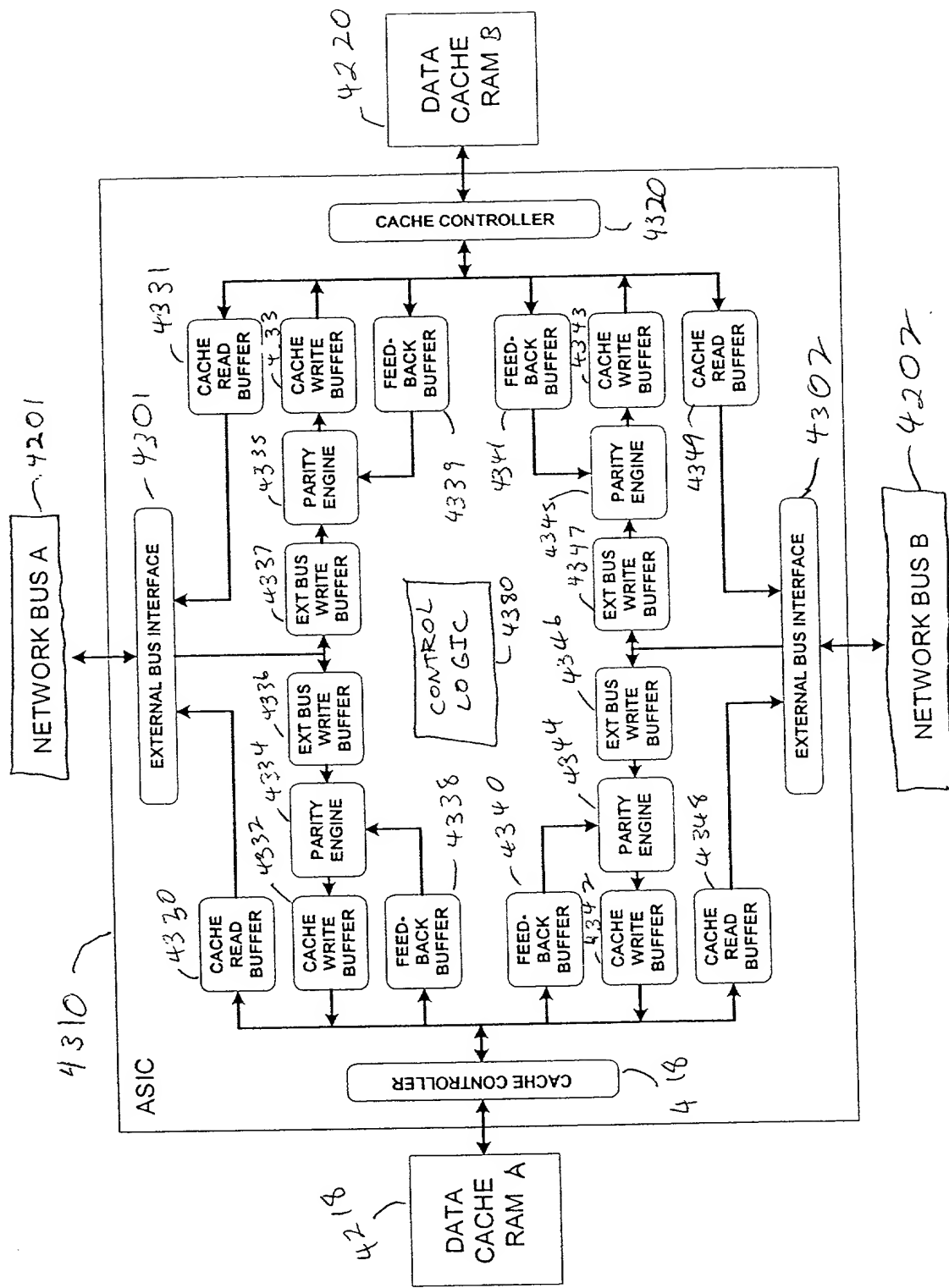


FIGURE 43

PCI map	Block Size	Opcode	Spare	Parity Index	Spare	RAM Adr
63-----62, 61	-----59, 58	-----56, 55	-----51, 50	-----35, 34, 32, 31	-----	-----0

4400

FIGURE 44